

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 19:46:25 ; Search time 132 Seconds  
(without alignments)  
6180.135 Million cell updates/sec

Title: US-09-970-851-1

Perfect score: 1470

Sequence: 1 atgtctcccgattcgatga.....acacggactgggggtattga 1470

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/2/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/2/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/2/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/2/ina/PTCUS COMB.seq:\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106.6	7.3	1572	4	US-09-354-138-125
2	100.2	6.8	1578	4	US-09-354-138-134
3	98.6	6.7	15894	1	US-08-348-891A-1
4	98.6	6.7	15894	1	US-08-905-817-1
5	69	4.7	15462	3	US-09-073-492-1
6	49.2	3.3	7218	1	US-08-232-463-14
7	37.2	2.5	834	4	US-09-621-976-2574
8	36.4	2.5	68750	3	US-09-335-409-1
9	36.4	2.5	68750	4	US-09-568-102-1
10	36.4	2.5	68750	4	US-09-567-969-1
11	36.4	2.5	68750	4	US-09-568-480-1
12	36.4	2.5	68750	4	US-09-568-486-1
13	36.4	2.5	68750	4	US-09-568-472-1
14	36.4	2.5	68750	4	US-09-567-899-1
15	35.8	2.4	505	4	US-09-621-976-15639
16	32.8	2.2	297	4	US-09-252-991A-12783
17	32.6	2.2	2135	4	US-08-933-711B-17
18	32.4	2.2	248	3	US-09-007-005-32
19	32.4	2.2	248	3	US-09-244-796-32
20	32.4	2.2	277	3	US-09-007-005-3
21	32.4	2.2	277	3	US-09-244-796-3
22	32	2.2	31208	4	US-09-852-067-3
23	31.8	2.2	276	4	US-09-313-294A-3427
24	31.8	2.2	289	3	US-09-007-005-17
25	31.8	2.2	289	3	US-09-244-796-17
26	31.8	2.2	291	4	US-09-313-294A-4473
27	31.8	2.2	324	4	US-09-313-294A-7082

28	31.8	2.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
29	31.8	2.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
30	31.6	2.1	71989	4	US-09-443-501A-2	Sequence 2, Appli
31	31.4	2.1	528	4	US-09-621-976-13023	Sequence 13023, A
32	31.4	2.1	4041	3	US-09-105-537-36	Sequence 36, Appl
33	31.4	2.1	6396	4	US-09-620-312D-226	Sequence 226, App
34	31.4	2.1	36778	3	US-09-105-537-5	Sequence 5, Appli
35	31.2	2.1	500	3	US-09-141-000-2	Sequence 2, Appli
36	31.2	2.1	1599	2	US-08-417-495-3	Sequence 3, Appli
37	31.2	2.1	1599	2	US-08-284-391B-3	Sequence 3, Appli
38	31.2	2.1	1599	3	US-09-218-950-3	Sequence 3, Appli
39	31.2	2.1	1599	5	PCT-US92-01785-3	Sequence 3, Appli
40	31.2	2.1	1599	5	PCT-US95-00454-3	Sequence 3, Appli
41	31.2	2.1	3854	2	US-08-720-484A-1	Sequence 1, Appli
42	31.2	2.1	3854	3	US-08-953-823A-1	Sequence 1, Appli
43	31.2	2.1	3854	4	US-09-398-239-1	Sequence 1, Appli
44	31.2	2.1	3854	4	US-09-560-876A-1	Sequence 1, Appli
45	31	2.1	834	4	US-09-621-976-2574	Sequence 2574, Ap

#### ALIGNMENTS

#### RESULT 1

US-09-354-138-125

; Sequence 125, Application US/09354138

; Patent No. 6309647

; GENERAL INFORMATION:

; APPLICANT: Paolletti, Enzo

; APPLICANT: Tartaglia, James

; APPLICANT: Taylor, Jill

; APPLICANT: Gettig, Russell

; TITLE OF INVENTION: POXVIRUS - CANINE DISTEMPER VIRUS (CDV)

; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE

; NUMBER OF SEQUENCES: 139

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford, P.C.

; STREET: 530 Fifth Avenue, 25th Floor

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/354,138

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/472,379

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: US 08/416,646

; FILING DATE: 05-APR-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/224,657

; FILING DATE: 16-APR-1994

; APPLICATION NUMBER: US 08/073,962

; FILING DATE: 08-JUN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/776,867

; FILING DATE: 23-OCT-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/621,614

; FILING DATE: 30-NOV-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/938,283

; FILING DATE: 31-AUG-1993

; PRIOR APPLICATION DATA:

	/	APPLICATION NUMBER:	US 08/105,483	
	/	FILING DATE:	12-AUG-1993	
	/	PRIOR APPLICATION DATA:		
	/	APPLICATION NUMBER:	US 07/847,951	
	/	FILING DATE:	06-MAR-1992	
	/	PRIOR APPLICATION DATA:		
	/	APPLICATION NUMBER:	US 07/713,967	
	/	FILING DATE:	11-JUN-1991	
	/	PRIOR APPLICATION DATA:		
	/	APPLICATION NUMBER:	US 07,666,056	
	/	FILING DATE:	07-MAR-1991	
	/	ATTORNEY/AGENT INFORMATION:		
	/	NAME:	Frommer, William S.	
	/	REGISTRATION NUMBER:	25,506	
	/	REFERENCE/DOCKET NUMBER:	454310-2860	
	/	TELECOMMUNICATION INFORMATION:		
	/	TELEPHONE:	(212) 840-3333	
	/	TELEFAX:	(212) 840-0712	
	/	INFORMATION FOR SEQ ID NO:	125:	
	/	SEQUENCE CHARACTERISTICS:		
	/	LENGTH:	1572 base pairs	
	/	TYPE:	nucleic acid	
	/	STRANDEDNESS:	single	
	/	TOPOLOGY:	linear	
	/	MOLECULE TYPE:	cDNA	
	/	US-09-354-138-125		
	/	Query Match	7.3%; Score 106.6; DB 4; Length 1572;	
	/	Best Local Similarity	46.0%; Pred. No. 5.3e-25;	
	/	Matches	361; Conservative 0; Mismatches 424; Indels 0; Gaps 0;	
QY		499	TCTATCCAGCCTCAGTGTGGGTCACAGTAGCGAAGGCCCATCACTGCATATGAGACACA	558
DB		505	TCCATCTTGCTCAAATTTGGATCCTGCTAGCTAAAGCGGTGAATCTCTGTAATCTGCA	564
QY		559	GATGAGTCGGAACAAGAAGAAATCAATAAGTATCATGTCAGCAGGCGAGTCCAGAAGAAG	618
DB		565	GCCGACTCGAGATGAGAGGTGATTAAGTATACCAGCAAGACGTTGTGTCGGAGAA	624
QY		619	TACATCTCACCTGTATGTCAGAGTGCATTAACACTCAATCAGACATTTCTCTGSCA	678
DB		625	TTTAGAATGAACAAAATCTGGCTTGTATTTGTTAGAAACAGGAATGCTGAGGACCTATCT	684
QY		679	GTCGCGATTTCTTAGTTAGTCAGAGCTTAAGAGAGGCGCAATACGCGAGTGGAGCTCC	738
DB		685	TTGAGGGATTCATGTTGGCGCTCATCTTGGACATCAACAGTCCCCAGGAACAAGCTT	744
QY		739	ACGTATTPAACACTTAGTAGGGATGTAGACTCATACATCAGGAACAACGAGCTTACTGCA	798
DB		745	AGAATTGCTGAATGATTTGTGATATAGATAAATCAATTTGTGGAAGCTGGGTAGCTAGT	804
QY		799	TTCTTCCTTACACTCAAAATGGAATTAATACCAAGACATCAGCCCTAGCACTCAGCAGC	858
DB		805	TTCACTCCTAACTCAAGTTTGGCATTTGAACATGATATCCGGCTCTGGGTGTCATGAG	864
QY		859	CTCACAGGCGATATCCAAAAAGATGAAGCAGCTCATCGTTTTATTCGGATGAAGGAGAA	918
DB		865	TTTTTCCGAGAAATTAACAACATTTGAATCCCTCATGATGCTATATCAACAGATGGGTGAA	924
QY		919	AATGCGCGCTCATGACATTTGCTAGTGCACAGTATCAGATGAGCTTTTCACCGGCTGAG	978
DB		925	ACAGCCCGTACATGGTTATCTTTGGAAAACCTGTTTCAAAAACAAATTTAGTCAGGGTCC	984
QY		979	TATGCACAGCTTTATTTCTTTTGGCATGGGCATCGCTCTTATAGATAAGGAACTGCG	1038
DB		985	TACCATTTGCTCTGGAGTTATGCTATGGGGTTGGTTGGTTGAACTTGAAACTCCATGGA	1044
QY		1039	AAATACCAATTCGCCAGAGACTTCATGAGACATCATTTCTGGAGACTCGGGGTGGAGTAT	1098
DB		1045	GGGTAAATTTTCGGTCGATCTTACTTTTGAACCCAGCTTACTTCAGACTCGGGCAAGAAATG	1104
QY		1099	GCTCAGGCTCAGGGGAGTAGCATCAACGAAGACATGCGTCTGAGCTTAAATCTAACCCCG	1158





Db 793 TAGCGGATTTCATGGTGGCTCTAATCCTGGATATCAAGAGAACACCGGAAACAAACCCA 852  
Qy 740 CGTATTACAACTTAGTAGGGATGTAGACTCATATCATAGAGAACACCGGACTTACTGTCAT 799  
Db 853 GGATTGCTGAAATGATGTGACATTTGATACATATATCGTAGAGGAGGATTTAGCCAGTT 912  
Qy 800 TCTTCCTTACACTCAATATGAATTAATACCAAGACATCAGCCCTAGCACTCAGCAGCC 859  
Db 913 TTATCTGACTATTAAGTTTGGGATAGAACTATGTATCTCTTGGACTGCATGAAT 972  
Qy 860 TCACAGCGCATATCCAAAGATGAAGCAGCTCATGGCTTTATATCGGATGAAGGGAGAAA 919  
Db 973 TTGCTGGTGGTATCCACACTTGAGTCTCTTGATGAACCTTTACAGCAAAATGGGGAAA 1032  
Qy 920 ATGCGCGGTACATGACATTTCTTTGGCCATGGCATGGCATCAGTCTTAGATAAAGGAACCTGGCA 1039  
Db 1033 CTGCAACCTTACATGGTAAACCTGGAGAACTCAATTCAGAACAAAGTTTCAGTGCAGGATCAT 1092  
Qy 980 ATGCACAGCTTTATCTTTTGGCCATGGCATGGCATCAGTCTTAGATAAAGGAACCTGGCA 1039  
Db 1093 ACCCTCTGCTCTGGAGCTATGCGCATGGGAGTAGAGTGGAACTTGAATCTCCATGGGAG 1152  
Qy 1040 AATACCAATTCGCGAGAGACTTCAATGAGCACATCAITCTCGAGACTCGGGGTGGAGTATG 1099  
Db 1153 GTTTGAACTTTGGCCGATCTTACTTTGATCCAGCATATTTAGATTAGGGCAAGAGATGG 1212  
Qy 1100 CTGAGGCTAGGGAGTAGCATCAACGAAGACATGGCTGTGAGCTTAAACTAACCCCGG 1159  
Db 1213 TAAGGAGGTGAGTGGAAAGGTGAGTTCCACATTTGGCATCTGAATCTGGTATCATCTGCGG 1272  
Qy 1160 CAGCAAGAGGGCGCTGGCAGCTGCTGCCAACGAGTGTCTGAGGAACTGGCAGCGTGG 1219  
Db 1273 AGGATGCAAGCTGTTTTCAGAGATTGCAATGCAATCTAGAGACAGATCAGTAGAG 1332  
Qy 1220 ATATTCCTACTCAACAGCGGGTCTCTCACTGGGCTCAGCGATGGAGG 1268  
Db 1333 CGGTTGGACCCAGCAAGCCCAAGTATCATTTCTACAGGTTGATCAAG 1381

RESULT 5

US-09-073-492-1  
; Sequence 1, Application US/09073492  
; Patent No. 6248578  
; GENERAL INFORMATION:  
; APPLICANT: Banerjee, Amiya K.  
; APPLICANT: Hoffman, Michael A.  
; TITLE OF INVENTION: Infectious Clone for Human Parainfluenza  
; TITLE OF INVENTION: Virus Type 3  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Calfee, Halter & Griswold LLP  
; STREET: 800 Superior Avenue  
; CITY: Cleveland  
; STATE: Ohio  
; COUNTRY: USA  
; ZIP: 44114-2688  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/073,492  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Doherly, Pamela A.  
; REGISTRATION NUMBER: 40,591  
; REFERENCE/DOCKET NUMBER: 23114/04000  
; TELEPHONE: (216)622-8200  
; TELEFAX: (216)241-0816  
; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15462 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-09-073-492-1

Query Match 4.7%; Score 69; DB 3; Length 15462;  
Best Local Similarity 53.5%; Pred. No. 9.5e-12;  
Matches 144; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

Qy 762 TGTAGACTCATATACATCAGGAACACCGGACTTACTGCTTCTTACACTCAATATGG 821  
Db 885 TGTGTGTAATACATAAGAGATGAGGTCTTGTCTTCAATACAAATCAGGTATGG 944  
Qy 822 AATTAATACCAAGACATCAGCCCTAGCAGCTCAGCAGCTCAGAGGAGATATCCAAAAGAT 881  
Db 945 AATTGAGCTAGATGGCAGCTTTGACTCTATCTCTCAGACCAGATATCAATAGATT 1004  
Qy 882 GAAGCAGCTCATCGCTTTATATCGGATGAAGGAGAGAAATGCGCCGTACATGACATTGCT 941  
Db 1005 AAAAGCTCTGATGCAATTTGATTTATCAAGGAGACCAAGGCTCTTTTATCTGTATCTCT 1064  
Qy 942 AGTGACAGTGTACAGATGAGCTTTGCACCGGCTGAGTATGACAGCTTTATCTTTTTC 1001  
Db 1065 CAGAGATCTTATCATGTTGAGTTCGACCCAGGCAACTATCTCTGCCATATGGAGTTATGC 1124  
Qy 1002 CATGGGATGGCATCAGCTCTTAGATAAAG 1030  
Db 1125 AATGGGGTGGCAGTTGTACAAAACAGAG 1153

RESULT 6

US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109

TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fls  
US-08-232-463-14

Query Match 3.3%; Score 49.2; DB 1; Length 7218;  
Best Local Similarity 7.7%; Pred. No. 2.5e-05;  
Matches 33; Conservative 210; Mismatches 183; Indels 0; Gaps 0;

QY 1023 AGATAAGGAAGTGGCAATACCAATTCGCCAGAGACATCATGAGACATCATTCGGAG 1082  
DB 1453 AGATAGAGATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1394  
QY 1083 ACTCGGGTGGAGTATGCTCAGGCTCAGGGAGTAGCATCAACGAAGACATGGCTGTGA 1142  
DB 1393 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1334  
QY 1143 GCTAAACATACCCCGGAGCAAGAGGGCTGCGAGCTGCTGCCAAGAGTGTCTGA 1202  
DB 1333 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1274  
QY 1203 GMAAACTGGCAGCGTGGATATTCCTACTCAACGAAGCGGGTCTCACTGGGCTCAGCA 1262  
DB 1273 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1214  
QY 1263 TGGAGGCCCGGAGCTCTCAGGGTGGATCGAACAAAGTCGAAGGCAACCATGCGCG 1322  
DB 1213 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1154  
QY 1323 AGATGGGAGACCAATTTCTGGATTGATGAGACAGTGGCGAAGACATGCGAGAAGC 1382  
DB 1153 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1094  
QY 1383 GCCAACTCCGACAGACACCAACCCAGCCCGGAGTCCCGGCGCATCA 1442  
DB 1093 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1034  
QY 1443 AGATAA 1448  
DB 1033 GAATTA 1028

RESULT 7  
US-09-621-976-2574  
; Sequence 2574, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Joibert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2574  
; LENGTH: 834  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 497..826  
US-09-621-976-2574

Query Match 2.5%; Score 37.2; DB 4; Length 834;  
Best Local Similarity 14.2%; Pred. No. 0.066;

Matches 27; Conservative 90; Mismatches 73; Indels 0; Gaps 0;

QY 55 ATGGAGCTACCGGAGGGGAGAGAGAGGAGACCTTTAAGAGTTGAGTCCAGTATTC 114  
DB 9 AAKWMSGSGRSSTRRKSSMGRSMCRGWSGCRSTRMRRSKSYGARRCMRRGKAKCM 68  
QY 115 ACTCTTAACAGTACGATCCAGAGATAGATCGAATTTTGGCGTATTCTGCTTCGGATT 174  
DB 69 SWTGMSCWSTGRMSSWRKRKRSWSWTWKWGRGYSWYISYRMYRCYSWMMWCMRSCYK 128  
QY 175 GCTGTTAGGAGACGCCCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGTGC 234  
DB 129 SRSRMMWMSGRSMCMYGSARWCWRYRMYRWTCAKSRMKCMWKKWSGSKCK 188  
QY 235 TCCGATTCTC 244  
DB 189 SSMCATGCTC 198

RESULT 8  
US-09-335-409-1  
; Sequence 1, Application US/09335409  
; Patent No. 6121029  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Goerlach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/335,409  
; CURRENT FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 68750  
; TYPE: DNA  
; ORGANISM: Sorangium cellulosum  
US-09-335-409-1

Query Match 2.5%; Score 36.4; DB 3; Length 68750;  
Best Local Similarity 48.1%; Pred. No. 2.1;  
Matches 103; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 1125 CGAAGACATGCTGCTGAGCTAAATAAATCCCGGCGAGCAAGAGGGCCCTGGCAGTGC 1184  
DB 31632 CGAGCAGCGGCTGACCCGTACCCCAACGAGCCCGCCGCAAGCATTTGCTGCCAGGCGC 31691  
QY 1185 TGCCCAACGAGTGTCTGAGGAACTGGCAGCGTGGATATTCCTACTCAACAGCGGGGT 1244  
DB 31692 TTTGCGAAGAGGGGTGTCTCCGCTCGACGTTGATTTTGTGGAGTGTACCGGACAGGGA 31751  
QY 1245 CCTCACTGGGCTCAGCGATGAGGCCCCCGGAGCCCTCTCAGGGTGGATCGAACAGTCCGA 1304  
DB 31752 CGGCGCTGGGCGACCCGATCGAGGTGCGAGGCGCTGAGCGAGGTGTATGTTCCAGGGCGCT 31811  
QY 1305 AGGCGAACCATGATCCGAGATGGGAGACCCAA 1338  
DB 31812 CCGGGGACCGACCGCTGTGTGGGGCGCGCAA 31845

RESULT 9  
US-09-568-102-1  
; Sequence 1, Application US/09568102  
; Patent No. 6346404  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon

APPLICANT: Goerlach, Joern  
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
FILE REFERENCE: 4-30582A  
CURRENT APPLICATION NUMBER: US/09/568,102  
CURRENT FILING DATE: 2000-05-10  
PRIOR APPLICATION NUMBER: 09/335,409  
PRIOR FILING DATE: 1999-06-17  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 68750  
TYPE: DNA  
ORGANISM: Sorangium cellulosum  
US-09-568-102-1

Query Match 2.5%; Score 36.4; DB 4; Length 68750;  
Best Local Similarity 48.1%; Pred. No. 2.1;  
Matches 103; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 1125 CGAAGCATGGCTGCTGAGCTAAACTAACCCCGGAGCAGAGGGGCTGCGAGCTGC 1184  
DB 31632 CGAGCAGCGGCTGACCGTACCCAAACCGACCCCGCAGCAAGCATTTGTCGCGCAGGCGC 31691

QY 1185 TGCCCAACAGTGTCTGAGGAACCTGCGAGGTGATATTCCTACTCAACAAGCCGGGT 1244  
DB 31692 TTTCGCAAGCAGCGGTGCTCCCGTTCAGCTTGAATTTGTGGAGTGTACGGGACAGGGA 31751

QY 1245 CTTCACTGGGCTCAGCGATGAGGCCCCCGAGGCTCTCAGGGTGGATCGAACAAGTCGCA 1304  
DB 31752 CGCGCTGGGCGACCGATCGAGTGCAGGCGCTGAGCGAGGTGTATGTTCCAGGGGCGCT 31811

QY 1305 AGGGCAACAGATGCCGAGATGGGAGAGCCCAA 1338  
DB 31812 CCGGGAGCCGACCGCTGCTGCTGGGGCGCGCAA 31845

RESULT 10  
US-09-567-969-1  
Sequence 1, Application US/09567969  
Patent No. 6355457  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James  
APPLICANT: Molnar, Istvan  
APPLICANT: Zirkle, Ross  
APPLICANT: Cyr, Devon  
APPLICANT: Goerlach, Joern  
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
FILE REFERENCE: 4-30582A  
CURRENT APPLICATION NUMBER: US/09/567,969  
CURRENT FILING DATE: 2000-05-10  
PRIOR APPLICATION NUMBER: 09/335,409  
PRIOR FILING DATE: 1999-06-17  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 68750  
TYPE: DNA  
ORGANISM: Sorangium cellulosum  
US-09-567-969-1

Query Match 2.5%; Score 36.4; DB 4; Length 68750;  
Best Local Similarity 48.1%; Pred. No. 2.1;  
Matches 103; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 1125 CGAAGCATGGCTGCTGAGCTAAACTAACCCCGGAGCAGAGGGGCTGCGAGCTGC 1184  
DB 31632 CGAGCAGCGGCTGACCGTACCCAAACCGACCCCGCAGCAAGCATTTGTCGCGCAGGCGC 31691

QY 1185 TGCCCAACAGTGTCTGAGGAACCTGCGAGGTGATATTCCTACTCAACAAGCCGGGT 1244  
DB 31692 TTTCGCAAGCAGCGGTGCTCCCGTTCAGCTTGAATTTGTGGAGTGTACGGGACAGGGA 31751

QY 1245 CTTCACTGGGCTCAGCGATGAGGCCCCCGAGGCTCTCAGGGTGGATCGAACAAGTCGCA 1304  
DB 31752 CGCGCTGGGCGACCGATCGAGTGCAGGCGCTGAGCGAGGTGTATGTTCCAGGGGCGCT 31811

QY 1305 AGGGCAACAGATGCCGAGATGGGAGAGCCCAA 1338  
DB 31812 CCGGGAGCCGACCGCTGCTGCTGGGGCGCGCAA 31845

RESULT 11  
US-09-568-480-1  
Sequence 1, Application US/09568480  
Patent No. 6355458  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James  
APPLICANT: Molnar, Istvan  
APPLICANT: Zirkle, Ross  
APPLICANT: Cyr, Devon  
APPLICANT: Goerlach, Joern  
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
FILE REFERENCE: 4-30582A  
CURRENT APPLICATION NUMBER: US/09/568,480  
CURRENT FILING DATE: 2000-05-10  
PRIOR APPLICATION NUMBER: 09/335,409  
PRIOR FILING DATE: 1999-06-17  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 68750  
TYPE: DNA  
ORGANISM: Sorangium cellulosum  
US-09-568-480-1

Query Match 2.5%; Score 36.4; DB 4; Length 68750;  
Best Local Similarity 48.1%; Pred. No. 2.1;  
Matches 103; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 1125 CGAAGCATGGCTGCTGAGCTAAACTAACCCCGGAGCAGAGGGGCTGCGAGCTGC 1184  
DB 31632 CGAGCAGCGGCTGACCGTACCCAAACCGACCCCGCAGCAAGCATTTGTCGCGCAGGCGC 31691

QY 1185 TGCCCAACAGTGTCTGAGGAACCTGCGAGGTGATATTCCTACTCAACAAGCCGGGT 1244  
DB 31692 TTTCGCAAGCAGCGGTGCTCCCGTTCAGCTTGAATTTGTGGAGTGTACGGGACAGGGA 31751

QY 1245 CTTCACTGGGCTCAGCGATGAGGCCCCCGAGGCTCTCAGGGTGGATCGAACAAGTCGCA 1304  
DB 31752 CGCGCTGGGCGACCGATCGAGTGCAGGCGCTGAGCGAGGTGTATGTTCCAGGGGCGCT 31811

QY 1305 AGGGCAACAGATGCCGAGATGGGAGAGCCCAA 1338  
DB 31812 CCGGGAGCCGACCGCTGCTGCTGGGGCGCGCAA 31845

RESULT 12  
US-09-568-486-1  
Sequence 1, Application US/09568486  
Patent No. 6355459  
GENERAL INFORMATION:  
APPLICANT: Schupp, Thomas  
APPLICANT: Ligon, James  
APPLICANT: Molnar, Istvan  
APPLICANT: Zirkle, Ross  
APPLICANT: Cyr, Devon  
APPLICANT: Goerlach, Joern  
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
FILE REFERENCE: 4-30582A  
CURRENT APPLICATION NUMBER: US/09/568,486  
CURRENT FILING DATE: 2000-05-10  
PRIOR APPLICATION NUMBER: 09/335,409  
PRIOR FILING DATE: 1999-06-17  
NUMBER OF SEQ ID NOS: 30





Db 238 WRARGGKRRAGKSSMMRSMRMSAGKARMCRRMMNSCREMSYSOMGSKCMSCRGTC 179  
QY 1292 CGAACAAAGTCGCAAGGGCAACCAAGATGCCGAGATGGGGAGACCCCAATTCTTGGATTGA 1351  
Db 178 AKWRYARYAKRYASSMGKYMGCRCWCYAKARMYGYRSTRSRTGSRGMKYRRRKMYYM 119  
QY 1352 TGAGAGCAGTGGCAACAGCATGGGAGAGCGCCAACTCCGCACAGAGCACCACCACC 1411  
Db 118 WKYMMWSWCCYRMGAAMYGMSARAYRMYASMSACMCSRMMKMMWSMWMRCWRSRYRCW 59  
QY 1412 CGGAACCCCCCCC 1424  
Db 58 MSGKWCYSCCGYC 46

Search completed: September 4, 2004, 00:24:42  
Job time : 139 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 18:08:25 ; Search time 862 Seconds  
(without alignments)  
7244.607 Million cell updates/sec

Title: US-09-970-851-1  
Perfect score: 1470  
Sequence: 1 atgtcttcgattcgatga.....acaccgactgggggtattga 1470

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_29Jan04:\*
- 1: Geneseqn1980s:\*
  - 2: Geneseqn1990s:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002s:\*
  - 7: Geneseqn2003as:\*
  - 8: Geneseqn2003bs:\*
  - 9: Geneseqn2003cs:\*
  - 10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1161.2	79.0	1747	AAC90470	NDV strai
2	1156.4	78.7	1801	ABK86576	Abk86576 Newcastle
3	1148.4	78.1	15186	Aaz44617	Aaz44617 Newcastle
4	1176.2	12.0	1885	Aaql4560	Aaql4560 Human par
5	166	11.3	15384	AAf55095	AAf55095 Nucleotid
6	166	11.3	15384	AAf55096	AAf55096 Nucleotid
7	161.2	11.0	15384	AAf55097	AAf55097 Nucleotid
8	159.6	10.9	15246	ABT13955	ABT13955 Simian Vi
9	152	10.3	1527	ABT13956	ABT13956 Human Cry
10	152	10.3	15246	ABT13954	ABT13954 Human Cry
11	109.2	7.4	1736	AAZ46837	AAZ46837 Morbilliv
12	106.6	7.3	1572	ABi98955	ABi98955 Canine di
13	106.6	7.3	15690	ABK15038	ABK15038 Canine di
14	106.6	7.3	15690	ABK15039	ABK15039 Canine di
15	106.6	7.2	18826	AAQ05123	AAQ05123 Nucleocap
16	105.6	7.0	1683	AAQ70424	AAQ70424 Sequence
17	103.4	7.0	1683	AAQ54673	AAQ54673 Rinder-pe
18	101.8	6.9	1684	AAV18270	AAV18270 Measles v
19	101.8	6.9	15894	AAV18265	AAV18265 Measles v
20	101.8	6.9	15894	AAV18273	AAV18273 HPIV-3 PR
21	101.8	6.9	15894	AAZ22907	AAZ22907 Nucleotid
22	101.8	6.9	15894	AAZ22902	AAZ22902 Nucleotid
23	100.2	6.8	1578	ABi98964	ABi98964 Measles v

24	100.2	6.8	15894	2	AAV18271	Measles v
25	100.2	6.8	15894	2	AAV18268	Measles v
26	100.2	6.8	15894	2	AAV18269	Measles v
27	100.2	6.8	15894	2	AAV18264	Measles v
28	100.2	6.8	15894	2	AAZ22901	Nucleotid
29	100.2	6.8	15894	2	AAZ22908	Nucleotid
30	100.2	6.8	15894	2	AAZ22905	Nucleotid
31	100.2	6.8	15894	2	AAZ22906	Nucleotid
32	98.6	6.7	15894	2	AAQ40480	Attenuate
33	98.6	6.7	15894	2	AAV18266	Measles v
34	98.6	6.7	15894	2	AAZ22903	Nucleotid
35	97	6.6	15893	2	AAZ22904	Nucleotid
36	97	6.6	15894	2	AAV18267	Measles v
37	76	5.2	15456	4	AAC85008	Nucleotid
38	76	5.2	15456	4	AAD27765	Nucleotid
39	74.4	5.1	15456	4	AAC85009	Nucleotid
40	74.4	5.1	15456	6	AAD27766	Nucleotid
41	73.2	5.0	15600	7	ACF03818	Human par
42	72.6	4.9	15462	2	AAV18273	HPIV-3 PR
43	72.6	4.9	15462	2	AAV18274	HPIV-3 Ve
44	72.6	4.9	15462	2	AAV83561	Nucleotid
45	72.2	4.9	6843	2	AAV83538	PCR-gener

ALIGNMENTS

RESULT 1  
AAC90470  
ID AAC90470 standard; DNA; 1747 BP.

- XX AAC90470;
- AC AAC90470;
- XX 13-MAR-2001 (first entry)
- DT NDV strain Beaudette C NP gene.
- DE NDV
- XX NDV
- XX NDV, Newcastle disease virus; NP; nucleoprotein; virucide; vaccine;
- XX Beaudette C strain; strain BI; leader region; avian pathogen; ds.
- XX Newcastle disease virus.
- XX WO200067786-A1.
- XX 16-NOV-2000.
- XX 05-MAY-2000; 2000WO-US006700.
- XX 05-MAY-1999; 99US-0132597P.
- XX 16-DEC-1999; 99US-0171072P.
- XX (UYMA-) UNIV MARYLAND BALTIMORE.
- XX Samal SK;
- XX WPI; 2001-024828/03.

New vaccine for Newcastle disease, for treating or preventing Newcastle disease in avian host, comprising attenuated infection Newcastle disease virus.

Claim 11; Page; 50pp; English.

The present sequence is claimed in a specification relating to the production of novel Newcastle disease virus (NDV) strains from cDNAs and improved live attenuated Newcastle disease vaccines. The vaccines are useful for preventing or treating Newcastle disease in birds. The recombinant NDV may be used to express proteins of other avian pathogens and can be used as a vaccine vector. The new NDV vaccine, unlike previous NDV vaccines, is less likely to revert back to virulence, more stable, and completely pathogenic. Note: The present sequence is not displayed in the specification but a GenBank number is given

SQ Sequence 1747 BP; 493 A; 467 C; 437 G; 350 T; 0 U; 0 Other;

Query Match		79.0%; Score 1161.2; DB 4; Length 1747;
Best Local Similarity		86.9%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 193; Indels 0; Gaps 0;		
QY	1	ATGCTCTCCGATTCGATGAAATACGAGCAGCTCTCGCTGCTCAGACTCGCCCCAATGGA 60
DB	67	ATGCTCTCCGATTCGATGAAATACGAGCAGCTCTCGCTGCTCAGACTCGCCCCAATGGA 126
QY	61	GCTCAGCGAGGGGAGAGAGAGGACATTTAAGAGTTGAGTCCAGTCCAGTATTCATCTTT 120
DB	127	GCTCATGAGGAGGGGAAAGGGAGTACCTTTAAAGTAGACGTCCTCGGATTCATCTTT 186
QY	121	AACAGTACGATPCACAGATAGATGGAATTTTGGGTATTCCTGCTTCGGAATTCGTCTT 180
DB	187	AACAGTATGATCCACAGATAGTGGAACTTTGGGTATTCCTGCTTCGGAATTCGTCTT 246
QY	181	AGCGAGGACGCCAAACAAACCGCTCAGGCAAGTGTCTCTCATATCCCTCTGTGTCCCAT 240
DB	247	AGCGAAGATGCCAAACAAACCACTCAGGCAAGTGTCTCTCATATCTCTTTATGTCTCCAC 306
QY	241	TCTCAAGTGATGAGGAAACCATGCTTCGCTTCGAGGAAACAGAAATGAGCTACACTGACT 300
DB	307	TCACAGTGATGAGGAAACCATGCTTCGCTTCGAGGAAACAGAAATGAGCTACACTGACT 366
QY	301	GTTCTTGAGTGCATGGTTTTTACCAGACGCTGCTCAGTTTCAACAAACAGGAGTGGGTG 360
DB	367	GTGCTTGAGATGTGATGGCTTTGGCCAAACGCTATGCCAGTTTCAACAAATAGAGTGGAGTG 426
QY	361	TCTCAGGAGACACACAGAGATTCATGTGTATGACAGGCTCTCTCCCTCGGGCGTGCAGT 420
DB	427	TCTGAAGAGACACACAGATTCGCGATGTATGACAGGCTCTCTCCCTCGGGCGTGCAGT 486
QY	421	AACGGTACTCCGTTTCGTCACCGCTGGGTTTCAAGATGATGACACAGAGATATCACTGAT 480
DB	487	AATGCGACCCGTTTCGTCACAGCGCGGGCCGAGATGATGACACAGAGATATCACTGAT 546
QY	481	ACTCTGGAAGAATCTGTCTATCCAGCTCAGGTATGGGTCAAGTAGCGAAGCCCATG 540
DB	547	ACCTGGAGAGGATCTCTCTATCCAGGCCCAAGTATGGGTCAAGTAGCAGAAAGCCATG 606
QY	541	ACTGCAATACAGACAGAGATGATTCGGAACACAGAAATCAATAAGTACATGACGAA 600
DB	607	ACTCGGTATGAGATCGAGATGATTCAGAAACACAGACGAATCAATAAGTATATGACGAA 666
QY	601	GGCAGAGTCCAGAAAGATCATCTCCACCTGTATGACAGGAGTCAATTCACATCACA 660
DB	667	GGCAGGCTCCAAAGAAATACATCTCTACCCGTATGACAGGACACAAATCCCAACTCAG 726
QY	661	ATCAGACATTCCTCTGGCAGTCCGCAATTTCTTAGTTAGCGAGCTTAAGAGAGCCGCAAT 720
DB	727	ATCAGACATTCCTCTGGCAGTCCGCAATTTCTTAGTTAGCGAGCTTAAGAGAGCCGCAAC 786
QY	721	ACGGCAGGTGGGAGTCCAGTATTAACACTTAGTAGGAGGATGTAGACTCATACATCAGG 780
DB	787	ACGGCAGGTGGTACCTCTACTTTATTATACCTTAGTAGGAGGACGTAGACTCATATACAG 846
QY	781	AACACCGGACTTACTGCATTCCTCTTACACTCAATATGGAATTAATACCAAGACATCA 840
DB	847	AATACCGGCTTACTGCATTCCTCTTACACTCAATATGGAATTAATACCAAGACATCA 906
QY	841	GCCTTAGCACTCAGCAGCTCAGCGGATATCCAAAGATGAAGCAGCTCATGCGCTTTA 900
DB	907	GCCCTTGCACTTAGTAGCTCTCAGCGGACATCCAGAAGATGAAGCAGCTCATGCGTTTG 966
QY	901	TATCGATCAAGGGAGAAATCGCGGTACATGACATTCGTAGGTGACAGTATCAGATG 960
DB	967	TATCGATCAAGGGAGAAATCGCGGTACATGACATTCGTAGGTGACAGTATCAGATG 1026
QY	961	AGCTTTCACCGGCTGAGTATGCACAGCTTTATCTTTTGGCATGGGCGTGCATCAGTNC 1020
DB	1027	AGCTTTCGCGCTGCCGATGATGCACAACTTTACTCTCTCGCCATGGGTATGGCATCAGTC 1086

QY	1021	TTAGATAAAGGAACCTGGCAAAATACCAATTCGCCAGAGACTTCATGAGCACATCATTTCTGG 1080
DB	1087	CTAGATAAAGGTACTGGGAAATACCAATTTGCCAGGACTTTTATGAGCACATCATTTCTGG 1146
QY	1081	AGACTCGGGTGGAGTATGCTCAGGCTCAGGGAGTAGCATCAACGAGACATGGGTGCT 1140
DB	1147	AGACTTGGAGTAGAGTAGCTCAGGCTCAGGGAAGTAGCATTAACGAGGATATGGGTGCC 1206
QY	1141	GAGCTAAACTAACCCCGGACAGCAAGAGGGGCTTGGCAGCTGCTGCCCAACGAGTGTCT 1200
DB	1207	GAGCTAAAGTTAAACCCGACAGCAAGGAGAGGCTTGGCAGCTGCTGCCCAACGAGTGTCT 1266
QY	1201	GAGGAACTGGCAGCGTGGATATCTCTACTCAACAGCCGGGCTCTCTCACTGGGCTCAGC 1260
DB	1267	GAGGAGACAGCAGCATAGACATGCTACTCAACAGTCGGAGTCTCTCACTGGGCTCAGC 1326
QY	1261	GATGAGGCCCCCGAGGCTCTCAGGCTGGATCGAACAAGTTCGCAAGGGCAACCATGATGCC 1320
DB	1327	GAGGGGGTCCCAAGCCCTTCAAGCGGATCGAATAGATCGCAAGGGCAACCATGATGCC 1386
QY	1321	GGAGATGGGAGAGCCCAATTTCTTGATTTGATGAGCAGTGGCGAACAGCATGCGAGAA 1380
DB	1387	GGGATGGGAGAGCCCAATTTCTTGATTTGATGAGCAGTGGCGAACAGCATGCGAGAA 1446
QY	1381	GGCCAACTCCGACAGACACACCCACCCCGGAAACCCCGGACTCCCGGGCCATCA 1440
DB	1447	GGCCAACTCTGACAGAGGCACTCCCAATCGGGGCTCCCGCAACTCTCTGGGCCATCC 1506
QY	1441	CAAGATAACACACACGACTGGGGGTATTGA 1470
DB	1507	CAAGATAACACACGACTGGGGGTATTGA 1536

## RESULT 2

ABK86576  
ID ABK86576 standard; cdna; 1801 BP.

XX ABK86576;

AC AC  
XX 24-SEP-2002 (first entry)

DT Newcastle disease virus nucleoprotein cDNA.

DE Nucleoprotein; NP; immunodominant epitope; newcastle disease; poultry;  
KW virucide; ND; vaccine; ovo administration; ss; gene.

XX Newcastle disease virus.

PH Key Location/Qualifiers  
FT CDS 122..1588  
FT /\*tag= a  
FT /product= "Nucleoprotein"

XX MO200236617-A2.

XX 10-MAY-2002.

XX 30-OCT-2001; 2001WO-EP012573.

XX 02-NOV-2000; 2000EP-00203814.

XX (ALKU ) AKZO NOBEL NV.

XX Mebatsion T, Koolen MJM;

XX WPI; 2002-527441/56.

XX P-PSDB; AAU99172.

XX New Newcastle disease virus (NDV) mutant, useful as a vaccine for  
PT protecting poultry (e.g. chicken) against Newcastle disease or as a  
PT marker vaccine to distinguish vaccinated animals from those infected with  
PT naturally occurring NDV.



XX Newcastle disease virus LaSota genomic DNA.  
 XX Avian-paramyxovirus; infection; lentogenic; F protein; vaccine;  
 KW respiratory disease; gastrointestinal disease; poultry pathogen;  
 KW local immunity; ds.  
 XX Newcastle disease virus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 122..1591  
 FT /\*tag= a  
 FT CDS 1887..3074  
 FT /\*tag= b  
 FT CDS 3290..4384  
 FT /\*tag= c  
 FT CDS 4544..6205  
 FT /\*tag= d  
 FT CDS 6412..8145  
 FT /\*tag= e  
 FT CDS 8381..11995  
 FT /\*tag= f  
 XX WO9966045-A1.  
 XX 23-DEC-1999.  
 XX 17-JUN-1999; 99WO-NL000377.  
 XX 19-JUN-1998; 98EP-00202054.  
 XX (DIEN-) STICHTING DIENST LANDBOUWKUNDIG ONDERZOE.  
 XX Peeters BPH, De Leeuw OS, Koch G, Gielkens ALJ;  
 XX WPI; 2000-106102/09.  
 XX New avian paramyxovirus cDNA, useful for production of vaccine against  
 PT Newcastle disease virus.  
 XX Disclosure; Fig 3; 115pp; English.  
 XX This invention describes a novel avian-paramyxovirus cDNA (I) which  
 CC comprises a nucleic acid sequence corresponding to the 5' terminal end of  
 CC the genome of avian-paramyxovirus allowing the generation of an  
 CC infectious copy of avian-paramyxovirus. The cell line is useful for the  
 CC production of infectious lentogenic NDV (Newcastle Disease virus) without  
 CC the addition of exogenous proteolytic activity. Also it is possible to  
 CC generate a stable transfected cell line that expresses the wild-type F  
 CC protein in the virus envelope therefore providing infectious particles,  
 CC useful in the form of a vaccine, especially against respiratory and/or  
 CC gastrointestinal diseases. NDV can be easily cultured to very high titers  
 CC in embryonated eggs. Mass culture of embryonated eggs is relatively  
 CC cheap. NDV vaccines are relatively stable and can be simply administered  
 CC by mass application methods e.g. drinking water or by spraying or by  
 CC aerosol formation. The natural route of infection is by the respiratory  
 CC and/or gastrointestinal tract which are the major routes of  
 CC infection of many other poultry pathogens. NDV can induce local immunity  
 CC despite the presence of circulating maternal antibody. This sequence  
 CC represents the NDV strain LaSota genome DNA  
 XX  
 SQ Sequence 15186 BP; 4431 A; 3544 C; 3462 G; 3749 T; 0 U; 0 Other;  
 Query Match 78.1%; Score 1148.4; DB 3; Length 15186;  
 Best Local Similarity 86.3%; Pred. No. 0;  
 Matches 1269; Conservative 0; Mismatches 201; Indels 0; Gaps 0;  
 QY 1 ATGCTTCGCTATTCGATGATACGAGCGCTCCCTCGCTGCTCAGACTGCCGCCAATGGA 60  
 DB 122 ATGCTTCGCTATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181  
 QY 61 GCTCAGGAGGGGGAGAGAGAGGAGGAGCACTTTAAGAGTTGAGTCCCACTTTCACCTTT 120  
 DB 122 ATGCTTCGCTATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181

DB 182 GCTCATGGAGGGGAGAAAAGGAGGTACCTTAAAGTAGACGTCCGGTATTACACTCTT 241  
 QY 121 AACAGTGACGATCCAGAGATAGATGGAATTTTGGCGTATTCTGTCTTCGGATTGCTGTT 180  
 DB 242 AACAGTGATGACCCAGAGATAGATGAGCTTTGTGTATTCTGCTCCGGATTGCTGTT 301  
 QY 181 AGCGAGGACGCCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGTGCTCCCAT 240  
 DB 302 AGCGAAGATGCCAACAAACCACTCAGGCAAGGTGCTCTCATATCTCTTTATGCTCCAC 361  
 QY 241 TCTCAAGTGATCAGGAACCATGTTGCCCTTGCAGGAAAACAGAAATGAGGCTACACTGAT 300  
 DB 362 TCACAGGTAATGAGGAACCATGTTGCCATTGCGAGGAAAACAGAAATGAGGCTACACTG 421  
 QY 301 GTTCTTGAGATCGATGGTTTTTACCAGCAGCGTCCCTCAGTTCAAACAAGAGGTGGGTG 360  
 DB 422 GTGCTTGAGATTGATGGCTTTTGCCAAACGCGACGCCCTCAGTTCAACATAGGAGTGG 481  
 QY 361 TCTGAGGAGAGACACAGAGATTCAATGATGATGAGAGGTCTCTCCCTCGGCGTGCAGT 420  
 DB 482 TCTGAAGAGAGACACAGAGATTGCGATGATGAGCAGGATCTCTCCCTCGGCGATG 541  
 QY 421 AACGCTACTCCGTTTCGTCACGCGCTGGGCTTGAAGATGATGACCAAGAGATATCACTGAT 480  
 DB 542 AACGGAACCCGTTTCGTCACAGCGCGGCGAGAGATGATGACCAAGAGATATCACTGAT 601  
 QY 481 ACTCTGAAAGAAATCTCTGTTATCCAGGCTCAGGTATGCGGTCAAGTAGCAAGGCCCATG 540  
 DB 602 ACCCTGGAGAGATCCCTCTCTATCCAGGCTCAAGTATGGGTCAAGTAGCAAGGCCCATG 661  
 QY 541 ACTGCATATGAGACAGCAGATGAGTCCGAAACAAGAGAAATCAATAAGTATATGACGAA 600  
 DB 662 ACTCGTATGAGACTGAGATGAGTCCGAAACAAGAGGGAATCAATAAGTATATGACGAA 721  
 QY 601 GGCAGAGTCCAGAGAAAGTACATCTCCACCCCTGATGCGAGGTGCAATTCACACTACA 660  
 DB 722 GGCAGGCTCCAAAGAAATACATCTCTACCCGTATGCGAGGAGCAAAATCCAACTCAG 781  
 QY 661 ATCAGACATTTCTTGGCAGTCCGCAATTTCTTAGTTCAGGAGCTTAAAGAGGCGCGCAAT 720  
 DB 782 ATCAGACAGTCTCTTGGCAGTCCGCAATTTCTTAGTTCAGGAGCTTAAAGAGGCGCGCAAT 841  
 QY 721 ACCGCGAGTGGGAGCTCCAGATTAACAATTAGTAGGGGATGAGACTCATACATCAGG 780  
 DB 842 ACCGCGAGTGGTACCTCTACTTATTAACTGCTAGGGGACCTAGACTCATACATCAGG 901  
 QY 781 AACACCGGACTTACTGCAATTTCTTACACTCAAATATGGAATTAATACCAAGACATCA 840  
 DB 902 AATACCGGCTTACTGCAATTTCTTGTGACTCAAGTACGGAATCAACCAAGACATCA 961  
 QY 841 GCCTTAGCACTCAGCAGCCTTCAAGCGGATATCCAAAGATGAAGCAGCTCATGCCGTTTA 900  
 DB 962 GCCTTTGCACTTAGTAGCCTCTCAGCGGACATCCAGAGATGAAGCAGCTCATGCCGTTT 1021  
 QY 901 TATCGATCAAGCGGAGAAATGCGGTACATGACATTCCTAGGTGACAGTGAATGATG 960  
 DB 1022 TATCGATCAAGCGGAGAAATGCGGTACATGACATTCCTAGGTGACAGTGAATGATG 1081  
 QY 961 AGCTTTGCAACCGGCTGAGTATGACAGCTTTATTTCTTTGCCATGGCATGGCATCAGTC 1020  
 DB 1082 AGCTTTGCGCTCGCGAGTATGCAACAATTTACTCTTTGCTTGGTATGGCATCAGTC 1141  
 QY 1021 TTAGTAAAGAACTGGGCAATPACCAATTCGCGAGAGCTTCATGAGCAGCATCTTCG 1080  
 DB 1142 CTAGATAAAGGTACTTGGGAAATPACCAATTTGCGGAGCTTTATGAGCAGCATCTTCG 1201  
 QY 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGAGTAGCATCAACGAGACATGGCTGT 1140  
 DB 1202 AGACTTGGAGTAGATGATGCTCAGGCTCAGGGAAGTAGCATTAACGAGGATATGGCTGC 1261  
 QY 1141 GAGCTAAATTAACCCCGGCGAGCAAGAGGGGCTTGGCAGCTGTGCCCCAAGAGTGTCT 1200  
 DB 1262 GAGCTAAAGCTAAACCCCGAGCAATGAAGGGGCTTGGCAGCTGTGCCCCAAGAGTGTCT 1321

QY 1201 GAGGAACTGCGAGCGTGGATATTCCTACTCAACAGCCGGGTCTCTACTCGGCTCAGC 1260  
 DB |||||  
 QY 1322 GATGATACAGCAGCATATACATGCTACTCAACAGTCCGAGTCTCTACTGGCTTAGC 1381  
 DB |||||  
 QY 1261 GATGAGGCCCGAGCCTCTCAGGTGATCGAACAGTCCGAGGGCAACAGATGCC 1320  
 DB |||||  
 QY 1382 GAGGGGGGTCGCAAGCTCTACAAGCGGATCGAATAGATCGCAGGGCAACAGAGCC 1441  
 DB |||||  
 QY 1321 GGAGATGGGGAGACCCAAATCTTGGATTTGATGAGAGCAGTGGCCAAACAGCATGCCAGAA 1380  
 DB |||||  
 QY 1442 GGAGATGGGGAGACCCAAATCTTGGATTTGATGAGAGCAGTGGCCAAACAGCATGCCAGAA 1501  
 DB |||||  
 QY 1381 GGGCCAAACTCGGCACAGAGCACACCCACCGGAAACCCCGGACTCCCGGGCCATCA 1440  
 DB |||||  
 QY 1502 GGGCCAAACTCGGCACAGAGCACCTCCCAATCGGGGCTCCCGCAACTCTCGGGCCATCC 1561  
 DB |||||  
 QY 1441 CAAAGTAAACGACCGGACTGGGGGTATTGA 1470  
 DB |||||  
 QY 1562 CAAAGTAAACGACCGGACTGGGGGTATTGA 1591  
 DB |||||

RESULT 4  
 AAQ14560  
 ID AAQ14560 standard; cdna; 1885 BP.  
 AC AAQ14560;  
 XX  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 24-FEB-1992 (first entry)  
 XX  
 XX Human para-influenza type 4A virus nucleocapsid protein gene.  
 DE  
 XX  
 XX Vaccine; antigen; ss.  
 XX  
 XX Human parainfluenza virus 4a.  
 OS  
 XX  
 PN JP03236783-A.  
 XX  
 PD 22-OCT-1991.  
 XX  
 PF 26-NOV-1990; 90JP-00317931.  
 XX  
 PR 27-DEC-1989; 89JP-00336291.  
 XX  
 XX (FUJG ) FUJIKURA CHEM CO LTD.  
 PA  
 XX  
 XX WPI; 1991-351134/48.  
 DR  
 XX  
 XX The gene can be used for the early diagnosis of human para-influenza type  
 CC 4A viral infection. Protein expressed from the gene can be used for the  
 CC prep. of antigens and vaccines. The sequence was obt. from a cDNA  
 CC generated from mRNA isolated from monkey kidney cells infected with the  
 CC virus. The position of the ORF is not given. (Updated on 25-MAR-2003 to  
 CC correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 FS Claim 1; Page 1; 9pp; Japanese.  
 XX  
 XX The gene can be used for the early diagnosis of human para-influenza type  
 CC 4A viral infection. Protein expressed from the gene can be used for the  
 CC prep. of antigens and vaccines. The sequence was obt. from a cDNA  
 CC generated from mRNA isolated from monkey kidney cells infected with the  
 CC virus. The position of the ORF is not given. (Updated on 25-MAR-2003 to  
 CC correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 1885 BP; 652 A; 394 C; 351 G; 488 T; 0 U; 0 Other;  
 Query Match 12.0%; Score 176.2; DB 2; Length 1885;  
 Best Local Similarity 48.0%; Pred. No. 3.7e-44;  
 Matches 567; Conservative 0; Mismatches 608; Indels 6; Gaps 2;  
 QY 1 ATGCTTCCGTATTCGATGAATACGAGCAGCTCTCGCTGCTCAGACTCGCCCGCAATGGA 60  
 DB |||||  
 QY 34 ATGCTTCCGTATTCGATGAATACGAGCAGCTCTCGCTGCTCAGACTCGCCCGCAATGGA 93  
 DB |||||

QY 61 GCTCAGCGGAGGAGAGAGGAGGAGCACTTTAAGAGTTGAGTCCAGTATTTACTCTT 120  
 DB |||||  
 QY 94 GGTGATCAACAATTTGTGCAATCTGATCTTTAAAGCAGAAATCCCTGTATTTGACTT 153  
 DB |||||  
 QY 121 AACAGTGCAGATCCAGAAAGATAGATGGAATTTTGGGGTATTTCTGCTTCGGATTTGCTT 180  
 DB |||||  
 QY 154 AATACAAATGATCCACAGCAAGATTCACCTAAATGAATTTTGTCTCCGGCAAGCTGTA 213  
 DB |||||  
 QY 181 AGGAGGAGCCCAACAAACCGCTCAGGCAAGGTCTCTCATATCCCTCTGCTGCTCCCAT 240  
 DB |||||  
 QY 214 TCATCATCTGCCAAATTCGGCAATTTAAACAGGGGCAATTAATATCTCTGCTTCCTTACAG 273  
 DB |||||  
 QY 241 TCTCAAGTCATCAGGAACCATGTTGCCCTTTCAGGHAACAGAAATCAGGCTACAGTACT 300  
 DB |||||  
 QY 274 GCCACATCAATGCAGAAATCATCTTATGATTGCTGCCAGAGCCAGATGCTGCTCTTAGG 333  
 DB |||||  
 QY 301 GTTCTTTGAGATCGATGGTTTTT-----ACCAGCAGCGTGCCTCAGTTTCAACAGAGAGTG 355  
 DB |||||  
 QY 334 ATCAITTGAAGTTGATGCAATAGACCCGCCAGATTTATACATTAATTAATTCAGAAATCT 393  
 DB |||||  
 QY 356 GGGTGTCTGAGGAGAGGACAGAGATT-CATGGTGTATAGCAGGGTCTCTCCCTCGGGCG 414  
 DB |||||  
 QY 394 GGCTCGGATGACATCAAGATCAGAGCTTATAGAGCATTTCCAGAGATTTGCCTATTAGC 453  
 DB |||||  
 QY 415 TGCAGTAAACGGTACTCCGTTTCGTCAAGGCTGGGGTTGAAGATGATGCCAGAGATATC 474  
 DB |||||  
 QY 454 TTAGCTGATAGGACAGTATTTGTATCCAGAGATGCAGAACATGCAGTATCGGATGATATG 513  
 DB |||||  
 QY 475 ACTGATATCTCGAAAGATTCCTGTCTATCCAGGCTCAGGTATGGGTACAGTAGCGAAG 534  
 DB |||||  
 QY 514 GATACATATCTGAATAGGATATTCAGTGTATTAATACAGGTTTGGATAATGGTGTGAAA 573  
 DB |||||  
 QY 535 GCATGATCTCATATGAGCAGCAGATGAGTCCGAAACAGAGAAATCAATTAAGTACATG 594  
 DB |||||  
 QY 574 TGCATGACAGCTTATGATCAACCCACTGGATCTGAGATCTGAGAGAGAGATTTAGCCAGTATAAG 633  
 DB |||||  
 QY 595 CAGCAAGGCGAGTCCAGAAAGATACATCTCCACCCCTGTATGACAGAGTGCAGTAAATCAA 654  
 DB |||||  
 QY 634 CAGCAGGAGTAAATGCTGGAGAGATATCAATGCAACAGATGCCAGAAATCATCCAA 693  
 DB |||||  
 QY 655 CTCATAATCAGACATCTCTGGCAGTCCGCAATTTTCTTAGTACGAGCTTTAGAGAGGC 714  
 DB |||||  
 QY 694 CTAGTAATAAGAGAAAGTATGGTTATAAGCAATTTCTTGTACAGGAAATGTTAACAGCT 753  
 DB |||||  
 QY 715 CGCAATACGCGAGTGGGAGCTCCACGTTATTAACAATTTAGTAGGGGATGTAAGTCTATAC 774  
 DB |||||  
 QY 754 GATAAGGTAGTGTCTTATACACAGGTTATATGCCATGTTGGGAGACATTTGCAAAATAT 813  
 DB |||||  
 QY 775 ATCAGGAACACCGGACTTTACTGCAATTTCTTCTTACACTCAAAATATGGAATTAATACCAAG 834  
 DB |||||  
 QY 814 ATAGCTAATGTCGGAATGAGCGCTTTCTTTCTTAGCTTAAAGTTTCGGCTTGGTAATAGA 873  
 DB |||||  
 QY 835 ACATCAGCCCTAGCACACTCAGCAGGCTCAGCGGATATCCAAAGATCAAGCAGCTCATG 894  
 DB |||||  
 QY 874 TGGAAACCACTTGTCTTTAGTGTGCAATTTCTCGGTGAACCTAGTTAACTGAAATCTCTTAGT 933  
 DB |||||  
 QY 895 GCTTTATATCGATGAAGGAGAAATGCGCGGTACATGACATTTGCTAGGTGACAGTGAAT 954  
 DB |||||  
 QY 934 TCATCATACAGAAAGTTAGGTGACAGATCAAGATATTTGGCTCTCTTGGAAATCTCTGAA 993  
 DB |||||  
 QY 955 CAGATGAGCTTTGCAACCGGCTGAGTATGACAGCTTTTATTTCTTTGCAATGGGATGGCA 1014  
 DB |||||  
 QY 994 CTGATGGAATTTGCTCCAGCTAAATACCCACTATTGTTTCAAGTTATGCTATGGAGTTGGA 1053  
 DB |||||  
 QY 1015 TCAGTCTTTAGATAAAGGAACCTGGCAATATCCAAATTCGCGAGAGACTTCATGAGCACATCA 1074  
 DB |||||  
 QY 1054 AGTGTTCAGATCCGCTGATCAGAAATTTATCAGTTTGGAGAGAACTTCTTAAATCCAGT 1113  
 DB |||||  
 QY 1075 TTCTCGAGACTCGGGGTGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAGACATG 1134  
 DB |||||  
 QY 1114 TACTTCCAGATGTTGTTGAGACTGCAATGAAACATCAAGGACAGTGTGATCTTAATTT 1173  
 DB |||||  
 QY 1135 GCTGCTGAGCTAAAACTAAACCCCGCAGCAGCAAGAGGGGCT 1175

Db 1174 GCCTCAGAACTTGAATACTGATGAACAGAGTGGACAT 1214  
 AAF55095  
 ID AAF55095 standard; DNA; 15384 BP.  
 XX AC AAF55095;  
 XX DT 15-MAY-2001 (first entry)  
 XX DE Nucleotide sequence of a Mumps virus genome.  
 XX KW Encapsidation protein; transcription protein; replication protein;  
 XX cell targeting; gene therapy; attenuated virus; vaccine; mumps; ss.  
 XX OS Mumps virus.  
 XX PN WO200109309-A2.  
 XX PD 08-FEB-2001.  
 XX PF 02-AUG-2000; 2000WO-US021192.  
 XX PR 02-AUG-1999; 99US-0146664P.  
 XX PR 23-JUN-2000; 2000US-0213654P.  
 XX PA (AMHP ) AMERICAN HOME PROD CORP.  
 XX PI Clarke DK, Johnson EJ, Sidhu MS, Udem SA;  
 XX WPI; 2001-123320/13.  
 PT Producing a recombinant mumps virus (MUV), useful as a mumps vaccine, by  
 PT transfecting or transforming a host cell with a transcription vector  
 PT comprising a MUV genome or antigenome, and an expression vector encoding  
 PT trans-acting proteins.  
 PS Claim 21; Page 95-100; 133pp; English.  
 CC The present sequence represents a Mumps virus genome. The specification  
 CC describes a method for producing a recombinant mumps virus. The method  
 CC comprises transfecting or transforming, in a rescue composition media, a  
 CC host cell with a transcription vector comprising a genome or antigenome  
 CC of mumps virus, and an expression vector encoding trans-acting proteins  
 CC (NP, P and L) necessary for encapsidation, transcription and replication.  
 CC The method is carried out under conditions sufficient to permit the co-  
 CC expression of the vectors and the production of the recombinant virus.  
 CC The recombinant virus has an ability to induce long-lasting immunity with  
 CC a single dose and a relatively low level of genome recombination. The  
 CC recombinantly produced Mumps viruses are useful in antibody generation,  
 CC diagnostic, prophylactic and therapeutic applications, cell targeting,  
 CC gene therapy, mutant virus preparation and immunogenic composition  
 CC preparation. The method may also produce an attenuated virus for use as a  
 CC vaccine for preventing or ameliorating mumps infection  
 XX SQ Sequence 15384 BP; 4758 A; 3356 C; 3136 G; 4134 T; 0 U; 0 Other;  
 Query Match 11.3%; Score 166; DB 4; Length 15384;  
 Best Local Similarity 47.5%; Pred. No. 2.1e-40;  
 Matches 558; Conservative 0; Mismatches 610; Indels 6; Gaps 2;  
 QY 1 ATGCTTCCGTATTCGATGAATACGACAGCTCTCGCTGCTCAGACTCGCCCAATGGA 60  
 Db 146 ATGTCATCTCTGCTCAAGGCAATTTGAGCGGTTACGATAGAACAGGAATTCAGACAGG 205  
 QY 61 GCTCACCAGGAGGAGAGAGGAGGACACTTTAGAGTTGAGTCCAGCTATTCACCTCTT 120  
 Db 206 GGTGAGAGGGTTCAATTCACCAGGAGCTTTAAAGTCAGCAGTCAAAGTCTTCGTTATT 265  
 QY 121 AACAGTGACGATCCAGAAAGATAGATGGAATTTTCGGGTATTTCTGCTTCGGATGCTGTT 180

Db 266 AACACACCCCAATCCACCAACACGCTATCAGATGCTAAACTTTTGTCTTAAGAATAATCTGC 325  
 QY 181 AGCGAGGACGCCAACAAACCGCTCAGGCAAGTGTCTCTCATATCCCTCTCTGTCTCCCAT 240  
 Db 326 AGTCAAAATGCTAGGCGATCTCACAGGCTAGGTGCATTGATAACATTTATCTCACTTCCC 385  
 QY 241 TCTCAAGTGTAGGAACCATGTTGCCCTTGAGGAAACAGAAATGAGGCTACACTGACT 300  
 Db 386 TCAGCAGGCATGCAAAATCATATTAGATTAGCAGATAGATCACCCGAAAGCTCAGATAGAA 445  
 QY 301 GTTCTTTGAGATCGATGGTTTT--ACCAGCAGCGTGCCTCAGTTTCAACAACAGAGTGGGG 358  
 Db 446 CGCTGTGAGATTGATGGTTTTTGAGCCTGTGATACATATAGGCTGATTTCCAATGACCGCC 505  
 QY 359 TGTCTGAGGAGAGACACAGAGATTCAT----GGTGATAGCAGGGTCTCTCCCTCGGGCG 414  
 Db 506 AATCTTTACTGCCAAATGAAATTTGCTGCTATGCTTTGCTTGCAGATGACCTCCCTCCACC 565  
 QY 415 TGCAGTAAACGCTACTCCGTTTCGTCACGCGTGGGTTGAAGATGATGCCACAGAGATATC 474  
 Db 566 ATAAATATGGAACCTCTTACGTACATGAGATGTTGAAGGACAGCCATGTGATGAGATT 625  
 QY 475 ACTGATACTCTGGAAGAAATCCTGTCTATCCAGGCTCAGGTATGGGTCAACAGTAGCGAG 534  
 Db 626 GAGCAGTTCTCTGATCGGTGTACAGTGTACTAATCCAGGCTTGGGTATGTGTCTGTA 685  
 QY 535 GCCATGACTGCTATGAGACAGCAGATGAGTCGGAACAAGAAAGATCAATAGTACATG 594  
 Db 686 TGTATGACAGCGTACGACCAACCTCGCGGTCTGTGATCGCGGATTTCCGAAATACACAG 745  
 QY 595 CAGCAAGGCAGAGTCCAGAAGATACATCTCCACCTCTGATGAGGAGTGAAGAGTCAATCAA 654  
 Db 746 CAGCAAGTCTGCTTGAGGCAAGATACATGTCACACCGGAGGCCCAAGGTGTGATTC 805  
 QY 655 CTCACATCAGACATTTCTTGCGAGTCCGCAATTTCTTAGTTAGCGAGCTTAAGAGAGCG 714  
 Db 806 ACTGCCATCAGAAAGTCTTGTGTTAGACAGTACCTTACCTTCAACTCCAGTCCAGTTGGG 865  
 QY 715 CGCATACGCGAGGTGGAGGCTCCAGTATTAACAATTTAGTAGGGGATGTAGACTCATAC 774  
 Db 866 AGACGGCAGGATTTGCTATCAACACAGATACATGCAATGTTGGGTGACATCGAAAGTAC 925  
 QY 775 ATCAGGAACACGGACTTTACTGCTTCTTCTTCACTCAATATGGAATTAATACCAAG 834  
 Db 926 ATTGAGAATTCAGGCTTACTGCTTCTTCTCACTCTCAAAATATGCACTAGGGACCAA 985  
 QY 835 ACATCAGCCTTAGCACTCAGCAGGCTCACAGCGATATCCAAAGATGAAGCAGCTCATG 894  
 Db 986 TGGAGTCTCTATATTGGCTGCAATTCACCGGTGAATCCACAGCTCCGATCCCTTGATG 1045  
 QY 895 CGTTTATTCGATGAAGGAGAAATGCGCGGTACATGACATTTGCTAGGTGACAGTGAAT 954  
 Db 1046 ATGTTATTCGAGTCTCGGAGAACAAAGCCAGATACCTTGTCTGTTAGAGGCTCCCCAA 1105  
 QY 955 CAGATGAGCTTTGACCCCGGGGCTACCCATTGATTTAGTTATGCTAGTTAGTGGAGTCCGT 1014  
 Db 1106 ATATGAGCTTTGACCCCGGGGCTACCCATTGATTTAGTTATGCTAGTTAGTGGAGTCCGT 1165  
 QY 1015 TCAGCTTTAGATAAGGAAATGCGCAATAACCAATTCGCCAGAGACTTTCATGAGCAGATCA 1074  
 Db 1166 ACAGTCTTAGATGTTCAAAATGCGAAATTAACACTTATGACGACCTTTCCTAAACGGTTAT 1225  
 QY 1075 TTCTGGAGACTCGGGGTGGAGTATGTCAGGCTCAGGGGAGTAGCATCAACGAAAGACATG 1134  
 Db 1226 TATTTCCAGATGGGGTTGAGACCGCACGGAAGCAACAAGGCACTGTTGCAACAGAGTA 1285  
 QY 1135 GCTGCTGAGCTAAACTAAACCCCGGAGCAAGAA 1168  
 Db 1286 GCAGATGATCTGGGCGCTGACTCTCTGACCAAGAA 1319

RESULT 6  
 AAF55096



II	AAAF5096 standard; DNA; 15384 BP.	QY	241	TCTCAAGTGATGAGGAACCATGTTGCCCTTGCGAGAAAACAGAAATGAGCGCTACACTGACT	300
XX	AAAF5096;	Db	386	TCAGCAGCATGCAAAATCATATTAGATTAGCAGATAGATCACCCGAAAGCTCAGATAGAA	445
XX	15-MAY-2001 (first entry)	QY	301	GTTCCTTGAGATCGATGGTTTT--ACAGCAGCGTGCCTCAGTTCAACAACAGAGTGGGG	358
XX	Nucleotide sequence of a Mumps virus genome.	Db	446	CGCTGTGAGATTGATGGTTTTGAGCCTGGTACATATAGGCTGATTCCTCAATGACCGCGC	505
XX	Encapsidation protein; transcription protein; replication protein;	QY	359	TGCTCTGAGGAGAGACAGACAGATTTCAT----GGTGATAGCAGGGTCTCTCCCTCGGGG	414
KW	cell targeting; gene therapy; attenuated virus; vaccine; mumps; ss.	Db	506	AATCTACTTGCCTGCAATGAAATTCCTGCTATGCTTGTCTTGCTTGAGATGACCTCCCTCCAACC	565
XX	Mumps virus.	QY	415	TGCAGTAACCGTACTCCGTTTCGTTCACGGCTGGGGTTGAAGATGATGACACAGAGATATC	474
XX	WO200109309-A2.	Db	566	ATAAATAATGGAATCTCTTACGTACATGAGATGTTGAAGGACAGCATGTGATGAGATT	625
XX	08-FEB-2001.	QY	475	ACTGATCTCTCGAAAGAAATCCTGTCTATCCAGGCTCAGGTATGGGTACAGTACAGGAG	534
XX	02-AUG-2000; 2000WO-US021192.	Db	626	GAGCAGTTCTCTGGATCGGTGTTTACAGTGTACTAATCCAGGCTTGGGTAAATGGTCTCTAA	685
XX	02-AUG-1999; 99US-0146664P.	QY	535	GCATGACTGTCATATGAGACAGCATGAGTCGGAACCAAGAAAGAAATCAATAAGTACATG	594
PR	23-JUN-2000; 2000US-0213654P.	Db	686	TGTATGACAGCGTACGACCAACCTGCCGGGTCTGCTGATCGGCGATTTGCGNAATACAG	745
XX	(AMHP ) AMERICAN HOME PROD CORP.	QY	595	CAGCAAGGCGAGAGTCCAGAGAAAGTACATCTCCACCCTGTATGCGAGGAGTGAATTCAA	654
XX	Clarke DK, Johnson EJ, Sidhu MS, Udem SA;	Db	746	CAGCAAGGTCGCTTTCAGGCAAGATACATGTCGAACCGGAGGCCCAAGGTTGATTCAA	805
XX	WPI; 2001-123320/13.	QY	655	CTCACAATCAGACATTTCTTGGCAGTCCGATTTCTTCTAGTTAGCGAGCTTAAGAGAGGC	714
XX	Producing a recombinant mumps virus (MUV), useful as a mumps vaccine, by	Db	806	ACTGCCATCAGGAAAAAGTCTTGTGTTAGACAGTACTTTACCTTCGAACTCCAGTTGGCG	865
PT	transferring or transforming a host cell with a transcription vector	QY	715	CGCAATACGCGAGTGGGAGCTCCAGTATTACAACTTAGTAGGGATGTAGACTCATAC	774
PT	comprising a MUV genome or antigenome, and an expression vector encoding	Db	866	AGACGCGCGAGGATGTTGTTATCAAAACAGATCTATGCAATGTTGGTGCATCGGAAGTAC	925
XX	trans-acting proteins.	QY	775	ATCAGGAACACCGGACTTACTGCAATTTCTCTTACACTCAAAATATGGAATTAATACCAAG	834
PS	Claim 21; Page 122-128; 133pp; English.	Db	926	ATTGAGNAATTCAGCGCTTACTGCTCTTCTTCTCACTCTCAAAATATGCACTAGGACCAAA	985
XX	The present sequence represents a Mumps virus genome. The specification	QY	835	ACATCAGCCTTAGCAGCTCAGCAGCCTCAGGGGATATCCAAAGATGAAGCAGCTCATG	894
XX	describes a method for producing a recombinant mumps virus. The method	Db	986	TGGAGTCTCTATCATTTGGCTGCAATTCACCGGTGAACCTCAACAGCTCCGATCCTTGATG	1045
CC	comprises transfecting or transforming, in a rescue composition media, a	QY	895	CGTTTATATCGATGAAGGAGAAATGCGCGTACATGACATGCTAGGTGACAGTGAT	954
CC	host cell with a transcription vector comprising a genome or antigenome	Db	1046	ATGTTATATCGAGTCTCGGAGAACCAAGCAGATACCTTGTCTGTAGAGCTCCCAAA	1105
CC	of mumps virus, and an expression vector encoding trans-acting proteins	QY	955	CAGATGAGCTTTGACCGGCTGAGTATGCAAGCTTTTATTCTTTTGGCATGGGATGGCA	1014
CC	(NP, P and L) necessary for encapsidation, transcription and replication.	Db	1106	ATAATGAGCTTTGACCGCGGGGCTACCCATTGATATTCAGTTATGCTATGGGAGTCGGT	1165
CC	The method is carried out under conditions sufficient to permit the co-	QY	1015	TCAGTCTTAGATAAAGAACTGGCAATATACAAATTCGCCAGAGACTTTCATGAGCACATCA	1074
CC	expression of the vectors and the production of the recombinant virus.	Db	1166	ACAGTCTCTAGATGTTCAAAATGCGAAATTTACACTTATGCAAGACCTTTCTTAAACGGTTAT	1225
CC	The recombinant virus has an ability to induce long-lasting immunity with	QY	1075	TTCTGGAGACTCGGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGGAAGACATG	1134
CC	a single dose and a relatively low level of genome recombination. The	Db	1226	TATTTCCAGATTGGGGTTGAGCCGCAAGAGACACAGGACATGTTTGACACAGAGTA	1285
CC	recombinantly produced Mumps viruses are useful in antibody generation,	QY	1135	GCTGCTGAGCTTAAAACTAAACCCCGGAGCAAGAA	1168
CC	diagnostic, prophylactic and therapeutic applications, cell targeting,	Db	1286	GCAGATGATCTGGGCTGACTCTCTGAGCAAGAA	1319
CC	gene therapy, mutant virus preparation and immunogenic composition	QY			
CC	preparation. The method may also produce an attenuated virus for use as a	Db			
CC	vaccine for preventing or ameliorating mumps infection	QY			
XX	Sequence 15384 BP; 4759 A; 3358 C; 3135 G; 4132 T; 0 U; 0 Other;	Db			
SQ	Query Match 11.3%; Score 166; DB 4; Length 15384;				
	Best Local Similarity 47.5%; Pred. No. 2.1e-40;				
	Matches 558; Conservative 0; Mismatches 610; Indels 6; Gaps 2;				
QY	1 ATGCTCTCCGTTATTCGATGAATACGAGCAGCTCTCGTGTGCTCAGACTCGCCCAATGGA	60			
Db	146 ATGTCATCTGTGCTCAAGGCATTTGAGCGGTTACGATAGAACAGAACTTCAAGACAGG	205			
QY	61 GCTACGAGGGGGAGAGAGGGAGCAGCTTTAAGAGTTGAGGTCCCAGTATTCATCTTT	120			
Db	206 GGTGAGGAGGGTTCAATTCACCGGAGACTTTAAAGTCAGCAGTCAAAAGTCTTCGTTATT	265			
QY	121 AACAGTACGATCCAGAGATAGATGGAATTTTGGGGTATTCTGTCTTCGGATTGCTGTT	180			
Db	266 AACACACCAATCCCAACACAGCTATCAGATGCTAAACTTTTGTCTTAAGATAATATGCG	325			
QY	181 AGCAGGAGCGCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCTCTCTGCTCCCAT	240			
Db	326 AGTCAAAATGCTAGGGCATCTCACAGGGTAGGTGCAATTGATAACATATTCTCACTTCCC	385			
XX	RESULT 7				
XX	AAAF5097				
XX	ID AAF55097 standard; DNA; 15384 BP.				
XX	AAAF5097;				
XX	15-MAY-2001 (first entry)				
XX					



KW epileptiform disease; multiple sclerosis; chronic fatigue syndrome;  
 KW primary lymphadenopathy-associated illness; gene therapy; gene; ds.  
 XX Simian parainfluenza virus 5.  
 XX WO200277211-A2.  
 PN 03-OCT-2002.  
 XX 07-FEB-2002; 2002WO-US004117.  
 XX 07-FEB-2001; 2001US-0267253P.  
 PR (CRYP-) CRYPTIC AFFLICTIONS LLC.  
 XX Robbins SJ;  
 PI WPI; 2003-040586/03.  
 DR  
 XX New nucleic acid useful for diagnosing and treating idiopathic  
 XX neurological disorders, including epileptiform diseases, e.g. epilepsy,  
 PT and lymphadenopathy-associated illnesses, and in screening of potential  
 PT new antiviral drugs.  
 XX Disclosure; Fig 9; 262pp; English.  
 PS  
 CC The invention relates to an isolated nucleic acid comprising: contiguous  
 CC nucleotide positions 1-15246 of a sequence comprising 15246 bp fully  
 CC defined in the specification; a nucleotide sequence complementary to the  
 CC 15246bp sequence; or a Cryptovirus-specific fragment of the 15246bp or  
 CC its complement, at least about 5 nucleotides long. The nucleic acid is  
 CC useful in diagnosing and treating many idiopathic neurological,  
 CC neurodegenerative, neuropsychological and neuropsychiatric disorders,  
 CC including epileptiform diseases (e.g. epilepsy, multiple sclerosis,  
 CC chronic fatigue syndrome or subacute sclerosing panencephalitis) and  
 CC primary lymphadenopathy-associated illnesses, and in research and  
 CC development, including screening of potential new antiviral drugs. The  
 CC nucleic acid, protein and the viral particle are useful in manufacturing  
 CC a vaccine. The protein is also used in producing a Cryptovirus-specific  
 CC antibody. The antibody may also be used in manufacturing a medicament for  
 CC the treatment of Cryptovirus infections. The polynucleotides of the  
 CC invention can be used to treat disorders by gene therapy. This  
 CC polynucleotide sequence represents the coding DNA for the Simian virus 5  
 CC protein relating to the invention. (Updated on 27-OCT-2003 to standardise  
 CC OS field)  
 XX  
 SQ Sequence 15246 BP; 4732 A; 3289 C; 3155 G; 4070 T; 0 U; 0 Other;  
 Query Match 10.9%; Score 159.6; DB 7; Length 15246;  
 Best Local Similarity 47.1%; Pred. No. 2.2e-38;  
 Matches 560; Conservative 0; Mismatches 624; Indels 6; Gaps 2;  
 QY 1 ATGTCCTCCGTTATTCGATGATACGAGCAGCTCCTCGCTGCTCAGACTCGCCCAATGGA 60  
 DB 152 ATGTCATCCGTCCTTAAGACATATGAGCGATTCAGCTCCTCACTCAAGAATCGAATCAG 211  
 QY 61 GCTCAGCAGGGGGGAGAGAGGGGAGCACTTTAAGAGTTGAGGTCCTCCAGTATTCATCTTT 120  
 DB 212 AGTGAGGAGGTACATCCACCTACACACTAAACCGGTAATCAGGATATTTATCTA 271  
 QY 121 AACAGTCAGATCCAGAGATAGATGGAATTTTGGGATATTCCTGCTTCGGATTCCTGTT 180  
 DB 272 ACCTCTAATAACCCAGAGTAAGATCCCGGCTCTTCTATTCCTGCTACGAGTTGTCTC 331  
 QY 181 AGCAGGAGCGCCACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGCTCCCAT 240  
 DB 332 AGTAATGGTGCAGGATTTCCATCGCTTTTGAGACATTAATCAATGTTTCGTACCA 391  
 QY 241 TCTCAAGTATGATGAGGAACCATGTTGCCCTTTCGAGGAAACAGAAATGAGGCTACATGACT 300  
 DB 392 TCAGCCCAATGCTCAATCATGTCAATTAATAGCTGACCACTCAGCAAGCTGATATCGAA 451  
 QY 301 GTTCTTGATGATCGATGGTTT--ACCAGCAGCGTGCCTCAGTTCAACACAGAGTGGGG 358

DB 452 AGGGTAGAGATCGATGGCTTTGAGGAGGATCATTCGCTTAATCCCAATCGAGTTCA 511  
 QY 359 TGTCTGAGGAGAGACAGAGATTTCAT---GGTGATAGCAGGGTCTCTCCCTCGGGG 414  
 DB 512 GGTATGAGCCGTGGAGAGATCAATGCCTATGCTGCACTTGCAGAGAGNCTACTGCACA 571  
 QY 415 TGCAGTAACGGTACTCCGTTTCGTTCACGGCTGGGGTTGAAGATGATGACACCAAGAGATATC 474  
 DB 572 CTAAACCATGCAACACCTTTTCGTTGATTCGGAAGTCGAGGAACTGCATGGGATGAGATT 631  
 QY 475 ACTGATACTCTCGAAAGAAATCCTGCTATTCAGGCTCAGGTATGGTTCACAGTAGCGAAG 534  
 DB 632 GAGACTTTCTTAGATATGTTTACAGTGTCTTAATGCAGGCATGGATAGTACTTCCAAAG 691  
 QY 535 GCATGACTGCATATGAGACAGACAGATGATCGGAAACAGAGAAATCAATAAGATACATG 594  
 DB 692 TGCATGACTGCGCCAGACCAACCTGCTGTTCTATTGAGAAACGCTGCAGAAATATCGT 751  
 QY 595 CAGCAAGGAGAGTCCAGAAAGAGTACATCTCCAGCCTGTATGTCAGGAGTGCATTTCAA 654  
 DB 752 CAGCAAGGAGAGTCAACCCGAGATATCTCTGCAACCGAGGCTCGACGATAATATCCAG 811  
 QY 655 CTCACAAATCAGACATTCCTGCGAGTCCGCAATTTTCTTAGTTCAGGAGCTTAAGAGAGGC 714  
 DB 812 AATGTAATCCGGAAGGGAATGGTGGTCAGACATTTTCTCTACCTTTGAACTGCGAGCTTGGC 871  
 QY 715 CGCAATAGCGCAGGTGGAGCTCCAGTATTTACAACTTACTAGGGGATGTAGACTCATAC 774  
 DB 872 CGAGCACAAGACCTTGTATCAAAATAGGTATTTATGCTATGTTAGGGGATGTTGGAAAGTAT 931  
 QY 775 ATCAGGAACACCGGACTTACTGCACTTCTCTTACACTCAAAATATGGAATTAATACCAAG 834  
 DB 932 ATAGAGATTTGTGAAATGGGAGGCTTCTTTTTCACACTAAATATGCAATAGGAACATAGA 991  
 QY 835 ACATCAGCCCTAGCCTCAGCAGCCTCAAGGCGATATCCAAAGATGAAGCAGCTCATG 894  
 DB 992 TGGCCACACACTTCTTTAGCTGCACTTTTTCAGGAGAGCTAACAAAGCTAAAAGTCCCTCATG 1051  
 QY 895 CGTTTATATCGGTGAAGGAGAAATGCGCGTACATGACATTTGCTAGGTGACAGATGAT 954  
 DB 1052 GCATTTATACAGACCCCTTGGTGACAGCCCGATTTTGGCCCTATTGGAGTCAACCAT 1111  
 QY 955 CAGATGAGCTTTGACACCGGCTGAGTATGCACAGCTTTTATTTTTCCTATGGGATGGCA 1014  
 DB 1112 TTGATGATTTTGTGCGACAAACTACCCACTGCTATATAGCTATGCTATGGGAATAGGC 1171  
 QY 1015 TCAGTCTTAGATAAAGGAACTGGCAATACCAATTCGCCAGAGACTTCATCAGCACATCA 1074  
 DB 1172 TATGTTTAGATGTCAACATGAGGAATACGCTTTTCTCCAGATCATATGAACAGACA 1231  
 QY 1075 TTCTGGAGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAGACATG 1134  
 DB 1232 TATTTCGAATTTGGGAATGGAACCTGCAAGAAACACACAGGTTGCACTGATGAGGATG 1291  
 QY 1135 GCTGCTGAGCTAAACCTAACCCCGCAGCAAGAGGGCCCTGGCAGCTGC 1184  
 DB 1292 GCAGAGATCTCGGCTTAACCTCAAGCCGACGACCGAGATGGCAATAC 1341

## RESULT 9

ABT13956

ID ABT13956 standard; DNA; 1527 BP.

XX ABT13956;

XX AC ABT13956 (first entry)

XX 13-FEB-2003 (first entry)

XX Human Cryptovirus strain BBR DNA region SEQ ID No 3.

XX Neuroprotective; virucide; anticonvulsant; antiinflammatory; Cryptovirus;  
 KW idiopathic; neurological; neurodegenerative; neuropsychological; vaccine;  
 KW epilepsy; neuropsychiatric disorder; subacute sclerosing panencephalitis;



KW epileptiform disease; multiple sclerosis; chronic fatigue syndrome;  
 KW primary lymphadenopathy-associated illness; gene therapy; gene; ds.  
 OS Rubulavirus sp.  
 XX WO200277211-A2.  
 FN 03-OCT-2002.  
 PD 07-FEB-2002; 2002WO-US004117.  
 XX 07-FEB-2001; 2001US-0267253P.  
 XX (CRYP-) CRYPTIC AFFLICTIONS LLC.  
 PA Robbins SJ;  
 XX WPI; 2003-040586/03.  
 DR New nucleic acid useful for diagnosing and treating idiopathic  
 PT neurological disorders, including epileptiform diseases, e.g. epilepsy,  
 PT and lymphadenopathy-associated illnesses, and in screening of potential  
 PT new antiviral drugs.  
 XX Claim 1; Fig 9; 262pp; English.  
 PS The invention relates to an isolated nucleic acid comprising: contiguous  
 CC nucleotide positions 1-15246 of a sequence comprising 15246 bp fully  
 CC defined in the specification; a nucleotide sequence complementary to the  
 CC 15246bp sequence; or a Cryptovirus-specific fragment of the 15246bp or  
 CC its complement, at least about 5 nucleotides long. The nucleic acid is  
 CC useful in diagnosing and treating many idiopathic neurological,  
 CC neurodegenerative, neuropsychological and neuropsychiatric disorders,  
 CC including epileptiform diseases (e.g. epilepsy, multiple sclerosis,  
 CC chronic fatigue syndrome or subacute sclerosing panencephalitis) and  
 CC primary lymphadenopathy-associated illnesses, and in research and  
 CC development, including screening of potential new antiviral drugs. The  
 CC nucleic acid, protein and the viral particle are useful in manufacturing  
 CC a vaccine. The protein is also used in producing a Cryptovirus-specific  
 CC antibody. The antibody may also be used in manufacturing a medicament for  
 CC the treatment of Cryptovirus infections. The polynucleotides of the  
 CC invention can be used to treat disorders by gene therapy. This  
 CC polynucleotide sequence represents the coding DNA for the human  
 CC Cryptovirus strain BBR protein of the invention  
 XX  
 SQ Sequence 15246 BP; 4753 A; 3273 C; 3131 G; 4089 T; 0 U; 0 Other;  
 Query Match 10.3%; Score 152; DB 7; Length 15246;  
 Best Local Similarity 46.1%; Pred. No. 5.4e-36;  
 Matches 549; Conservative 0; Mismatches 635; Indels 6; Gaps 1;  
 QY 1 ATGCTTCCTCGATTCGATGAATACGAGCAGCTCCTCGCTGCTCAGACTCGCCCAATGGA 60  
 DB 152 ATGTCATCCGTGTTAAAGCATATAGAGATTACACTCACTCAAGAACTCAAGATCAG 211  
 QY 61 GCTCAGCAGGGGAGAGAGAGGAGCCTTTAAGATTGAGTTCGCGGATTCACCTTT 120  
 DB 212 AGTCAAGAAGGCAATCCCACTACACACTAAACCGGTTATCAGGGTATTGTACTA 271  
 QY 121 AACAGTACGATCAGAGATAGATGGAATTTGCGGTATCTGCTTCGATTCGTT 180  
 DB 272 ACCTCTAATACCCAGAGCTAAGATCCCGGCTTCTTCTATTCTGCTACGATTTGTTCTC 331  
 QY 181 AGCAGGAGCGCAACAAACCGCTCAGGCAAGTGTCTTCATATCCCTCTGCTGCCAT 240  
 DB 332 AGTAATGTGCAAGGATTCCTATCGCTTTGGAGCATTTACTTACAATGTTTCGTACCA 391  
 QY 241 TCTCAAGTATGAGGAACCATGTTGCCCTTCAGGAAACAGAAATGAGGCTACACTGACT 300  
 DB 392 TCAGCTACATGCTCAATCATGTCAAAATTAGCTGACCAAGTCAACCAAGAGCTGATCGAA 451  
 QY 301 GTTCTTGATGCTGTTTACAGCAGCGCTGCTCAGTTCAACACAGGAGTGGGTG 360

DB 452 AGGGTAGAGATCGATGGCTTTTGGAGGGGATCAATTCGCTTAATTCCTCAATGCTCGCTCA 511  
 QY 361 TCTGAGGAGAGCACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 414  
 DB 512 GGTATGAGCCGTTGGAGAGATCAATGCCCTATGCTGCACTTTCGAGAGATCTACTGACACA 571  
 QY 415 TGCAGTAACGGTATCTCCGTTTCGTCACGGCTGGGGTTGAAGATGATGATGATGATGATGATG 474  
 DB 572 CTAAACCATGCAACACCTTTTCGTTGATTCGGAAGTCGAGGAACTGCTGAGGACGAGATT 631  
 QY 475 ACTGATCTCTGGAAGAAATCTGCTATCCAGGCTCAGGTATGGTACAGTACAGTACGGAAG 534  
 DB 632 GAGACTTTCTTAGATATGTTTACAGTGTCTTAATTCAGGCTGATGATGATGATGATGATG 591  
 QY 535 GCATGACTGATATGAGACAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 594  
 DB 692 TGCATGACTGCGCCAGACCAACCTGCTGCTTCTTATTTGAGAAACGCTGCAAAAATATCGT 751  
 QY 595 CAGCAAGGAGAGTCCAGAAAGATGATCATCTCCACCTGATGATGATGATGATGATGATGATGAT 654  
 DB 752 CAGCAAGGAGAGTCAACCCAGATATCTCTGCAACCGAGGCTGCAAGAAATATCCAG 811  
 QY 655 CTCACAAATCAGACATCTCTGCGAGTCCCATTTTCTTAGTTAGTACGAGCTTAAAGAGGC 714  
 DB 812 AATGTAATCCGGAAGGGAATGGTGGTTCAGACATTTCTCCTCACTTTGAACTGGAGCTTGGC 871  
 QY 715 CGCAATACGGCAGTGGAGCTCCAGTATTAACAATCTAGTAGGGGATGTAGACTCATAC 774  
 DB 872 CGAGCAAAAGCCTTGTATCAATAGTATATGCTATGTTAGGGGATGTTGGAAAGTAT 931  
 QY 775 ATCAGGAACACCGGACTTACTGCAATTTCTCTTACACTCAATATGGAATTAATACCAAG 834  
 DB 932 ATAGAGAAATTTGGAATGGAGGCTTCTTTTGGACACTAAATATGCAATTAGGAACCA 991  
 QY 835 ACATAGCCTTAGACTCAGACCTCAGGCGATATCAAGAGATGAGCAGCTCATG 894  
 DB 992 TGGCCACACATTTGCTTTAGCTGCAATTTTCAGGAGAGCTAACAAAGCTAAAGTCCCTCATG 1051  
 QY 895 CGTTATATCGGATGAAGGAGAAATGCGCCGTACATGACATTCGCTAGGTGACAGTGAT 954  
 DB 1052 GCATTGTACAGACCTTGTGTAGAGGCGCCGATATTTGGCCCTATTGGAGTACCACAC 1111  
 QY 955 CAGATGAGCTTTGACCGGCTGAGTATGACAGCTTTATTTTTCCTGTCATGGGCGATGGCA 1014  
 DB 1112 TTGATGATTTTGTGTCAGCAAACTATCCATTTGCTATATAGCTATGCTATGGGAATAGGC 1171  
 QY 1015 TCAGTCTTAGATAAAGAACTGGCAATACCAATTCGCGAGAGACTTTCATGAGCACATCA 1074  
 DB 1172 TATGTTTAGATGTCAACATGAGGAACCTACGCTTTCTCCAGATCATACATGAATAAGACA 1231  
 QY 1075 TTCTGGAGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAAGACATG 1134  
 DB 1232 TATTTCCAAATGGGAATGGAACTGCAAGAAACACAGGGTGCAGTTGACATGAGGATG 1291  
 QY 1135 GCTGCTGAGCTTAAACCTAACCCCGGAGCAAGAGAGGGCCCTGGCAGCTGC 1184  
 DB 1292 GCAGAAGATCTCGGTCTAATCTCAAGCCGACGACCCGAGATGGCAATATC 1341

RESULT 11  
 AAZ46837  
 ID AAZ46837 standard; cDNA to mRNA; 1736 BP.  
 XX AAZ46837;  
 AC  
 XX 06-AUG-2003 (revised)  
 DT 10-APR-2000 (first entry)  
 XX Morbillivirus antigenic protein encoding nucleotide sequence.  
 DE Antigenic; nucleocapsid protein; canine distemper virus; CDV;  
 KW Morbillivirus; Paramyxovirus; ss.  
 XX

OS Canine distemper virus.

PN JP11346768-A.  
 PD 21-DEC-1999.  
 XX  
 PF 03-JUN-1998; 98JP-00155072.  
 XX  
 PR 03-JUN-1998; 98JP-00155072.  
 XX  
 PA (MORG ) MORINAGA MILK IND CO LTD.  
 XX  
 DR WPI; 2000-109685/10.  
 DR P-PSDB; AAY56856.  
 XX

PT A protein having antigenicity of canine distemper virus nucleocapsid  
 XX protein - useful in the diagnosis of canine distemper virus infection.

PS Disclosure; Page 10-11; 15pp; Japanese.

XX The invention provides a protein which has a molecular weight of 58 kD  
 CC and has antigenicity of nucleocapsid protein (NP) of canine distemper  
 CC virus (CDV) belonging to Morbillivirus genus of Paramyxovirus family. The  
 CC protein can be used to prepare a reagent for the determination of anti-  
 CC CDV NP antibody by immobilizing the above protein on a carrier or a  
 CC membrane as the active component. The reagent can be used for diagnosis  
 CC of CDV infection, judgement of effect of vaccine and judgement of  
 CC inoculation period. The reagent can determine anti-CDV NP antibody  
 CC contained in canine serum easily. The present sequence represents a  
 CC nucleotide sequence encoding the Morbillivirus antigenic protein.  
 CC (Updated on 06-AUG-2003 to correct OS field.)  
 XX

SQ Sequence 1736 BP; 523 A; 386 C; 401 G; 426 T; 0 U; 0 Other;

Query Match 7.4%; Score 109.2; DB 3; Length 1736;  
 Best Local Similarity 47.5%; Pred. No. 4.6e-23;  
 Matches 324; Conservative 0; Mismatches 359; Indels 0; Gaps 0;

QY 499 TCTATCCAGGCTCAGGTATGGTACAGTACGAGGCGCATGCTGCATATGACAGCA 558  
 DB 510 TCCATCTGGCTCAAAATTGGATCTGCTAGCTAAGCGGTGATGCTCTGATGCA 569  
 QY 559 GATGAGTCGGAACAGAGATCAATAAGTACATGACGAGGCGAGTCCAGAAGAAG 618  
 DB 570 GCCGACTCGGAGATGAGAAGTGCTAAGTATATACCCAGCAAGACGTGTGTCGGAA 629  
 QY 619 TACATCTCCACCTGTATGACGAGTGAATTCACACTCAATTCAGACATCTCTGGCA 678  
 DB 630 TTAGAATGAACAAATCTGGCTGATATGTTAGAAACAGGATTGCTGAGGACCTATCT 689  
 QY 679 GTCCGCAATTTCTTAGTAGCGAGCTTAAGAGAGCGCCCAATACGCGAGGTGGAGCTCC 738  
 DB 690 TTGAGGCGATTCATGTTGGCACTCATTTTGACATCAAAAGATCCCAGGGAACACCT 749  
 QY 739 ACGTATTACACTTAGTAGGGATGTAGACTCATATCAGGAACACCGGACTTACTGCA 798  
 DB 750 AGAATTCTGAAATGATTTGTGTATATAGATAAATACATTGTGGAAGCTGGTGTAGT 809  
 QY 799 TTCTTCTTACACTCAAAATATGAATTAATACCAAGACATCAGCCCTAGCACTCAGCAGC 858  
 DB 810 TTCACTCAACTATCAAGTTGGCATTTGAACATATGATATCCGGCTCTTGGTGTGCATGAG 869  
 QY 859 CTCACAGGCGATATCAAAAGATGAAGCAGCTCATGCTTATATCGATGAAGGAGAA 918  
 DB 870 TTTTCGGGAAATTAACAACTATTGAATCCCTCATGATGCTATATCAACAGATGGGTGA 929  
 QY 919 AATGCGCGTACATGACATGCTTAGTGACAGTATGATGATGATGATGATGATGATGAT 978  
 DB 930 ACAGCACCATACATGTTATCTTTGGAATCTGTTCAAAACAAATTTAGTCAGGCTCC 989  
 QY 979 TATGCAACAGCTTATTTCTTTTGGCATGGGCGATCGCATCTCTAGTAAAGAACTGGC 1038  
 DB 990 TACCCATTCTCGGAGTTATGCTATGGGGGTTGGTGTGAATTTAGAACTCCATGGGA 1049

QY 1039 AAATACCAATTCGCCAGAGACTTCTATGAGCAATCTTGGAGACTCGGGGTGGAGTAT 1098  
 DB 1050 GGATTAATTTTCGTGATCTTACTTTGACCCAGCCTACTTCAGACTCGGGCAAGAAATG 1109  
 QY 1099 GCTCAGGCTCAGGGGAGTAGCATCAACGAGACATGGCTGCTGAGCTAAACTAACCCCG 1158  
 DB 1110 GTTAGACGATCTGCGAGGCAAGTAAGCTCCGCACTTGCTGCCGAGCTTGGCATCACAAG 1169  
 QY 1159 GCAGCAAGAAGGGGCTTGGCAG 1180  
 DB 1170 GAGGAAGCTCAGCTGCTGTCAG 1191

RESULT 12

ABI98955

ID ABI98955 standard; DNA; 1572 BP.

XX AC ABI98955;

XX DT 18-FEB-2002 (first entry)

XX DE Canine distemper virus, CDV, N gene.

XX KW Virucide; vaccine; virus; virulence; canine distemper virus; CDV;

XX KW measles; dog; ds.

XX OS Canine distemper virus.

XX PN US6309647-B1.

XX PD 30-OCT-2001.

XX PF 15-JUL-1999; 99US-00354138.

XX PR 15-JUL-1999; 99US-00354138.

XX PA (AVET ) AVENTIS PASTEUR.

XX PI Paoletti E, Tartaglia J, Taylor J, Gettig R;

XX DR WPI; 2002-040232/05.

XX Novel virus, useful for inducing immune response in dog against CDV,  
 PT comprises the modified recombinant virus having attenuated virulence  
 PT comprising exogenous DNA sequences encoding antigens of canine distemper  
 PT virus (CDV) or measles virus.  
 XX Claim 1; Fig 18; 147pp; English.

XX The present invention relates to modified recombinant viruses, comprising

XX inactivated virus-encoded genetic functions so that the viruses have  
 CC attenuated virulence, yet retained efficiency. The viruses can contain  
 CC DNA encoding a canine distemper virus (CDV) antigen or measles M or N  
 CC antigen. The recombinant viruses are useful for inducing an antigenic or  
 CC immunological response in a dog or other carnivore against CDV. The  
 CC present sequence was used in an example from the present invention

SQ Sequence 1572 BP; 472 A; 330 C; 372 G; 398 T; 0 U; 0 Other;

Query Match 7.3%; Score 106.6; DB 6; Length 1572;

Best Local Similarity 46.0%; Pred. No. 2.9e-22;

Matches 361; Conservative 0; Mismatches 424; Indels 0; Gaps 0;

QY 499 TCTATCCAGGCTCAGGTATGGTACAGTACGAGGCGCCATGCTCATATGACAGCA 558  
 DB 505 TCCATCTGGCTCAAAATTGGATCTGCTAGCTAAGCGGTGATGCTCTGATGCA 564  
 QY 559 GATGAGTCGGAACAGAGATCAATAAGTACATGCAAGGCGAGTCCAGAAGAAG 618  
 DB 565 GCCGACTCGGAGATGAGAAGGTGGATTAAGTATACCCAGCAAGAGCTGTGTCGGAA 624  
 QY 619 TACATCTCCACCTGTATGAGGAGTGAATTAACATCAGATCAGACATCTCTGGCA 678



Db 625 TTATGAAATGACAAATCTGGCTGATATTTGTTAGAAACAGGATTCCTGAGGACCTATCT 684  
 Qy 679 GTCCGCATTTTCTTAGTTAGCGAGCTTTAAGAGAGCGCGCAATACGGCAGGTGGAGCTCC 738  
 Db 685 TTGAGGCGATTCATGTTGGCGCTCATCTTGACATCAACGATCCCGCAGGAACAAGCCT 744  
 Qy 739 ACATATTACAACTTAGTAGGGATGTAGATCATATACATCAGGAACACCGGACTTACTGCA 798  
 Db 745 AGAATTCCTGAAATGATTTGTGATATAGATAAATACATTTGTGGAAGCTGGGTAGCTAGT 804  
 Qy 799 TTCTTCTTACACTCAAAATATGGAATTAATACCAAGACATCAGCCCTAGCACTCAGCAGC 858  
 Db 805 TTATCTTAACTATCAAGTTTGGATTTGAACTATGATTCGGCTCTTGGGTTCATGAG 864  
 Qy 859 CTCACAGCGGATATCCAAAAGATGAAGCAGCTCATGCGTTTATATCGGATGAAGGGAGAA 918  
 Db 865 TTTTCCGGAGATTTAAACAATTTGAAATCCCTCATGATGCTATATCAACAGATGGGTGAA 924  
 Qy 919 AATGCGCGTACATGACATTTGCTAGGTGACAGTGCATCAGATGACTTTGCACCGCTGAG 978  
 Db 925 ACAGCACCGTACATGTTATCTTGGAAACTCTGTTCAAAACAATTTAGTGCAGGCTCC 984  
 Qy 979 TATGCACAGCTTTATCTTTGGCATGGGATGGCATCAGTCTTAGATAAAGGAACCTGGC 1038  
 Db 985 TACCATTGCTCTGGAGTTATGCTATGGGGTGGTGTGAACTTGAAATCTCATCGGA 1044  
 Qy 1039 AAATACCAATTCGCGAGAGCTTCATGAGCACATCATTTGAGACACTCGGGGTGGAGTAT 1098  
 Db 1045 GGGTTAAATTTGGTGCATCTTACTTTGACCCAGCTTACTTCAGACTCGGGCAAGAAATG 1104  
 Qy 1099 GCTCAGGCTCAGGGAGTAGCATCAACGAAGACATGCTGCTGAGCTTAAACTAACCCCG 1158  
 Db 1105 GTTAGGAGATCTGCGGCAAGTAAGCTCTGCACCTTGCCTGCGGAGCTTGGCATCACCAAG 1164  
 Qy 1159 GCAGCAAGAGGGCGCTGGCAGCTGCTGCCCAACGAGTGTCTGAGGAACCTGGCAGCGTG 1218  
 Db 1165 GAGGAGCTCAGCTAGTGTGAAATATGATCCAGACACAGAGGACCGGACATTCGA 1224  
 Qy 1219 GATATTCCTACTCAACAGCGGGGTCTCTACTGGGCTCAGCGATGAGGCGCCCGGAGCC 1278  
 Db 1225 GCTACTGGTCTTAAGCAATCCAAATCACTTTTCTGCACTCGGAAGAGATCGGAAGTCGCC 1284  
 Qy 1279 TCTCA 1283  
 Db 1285 AATCA 1289

## RESULT 13

ABK15038 standard; cDNA; 15690 BP.

XX AC ABK15038;  
 XX DT 08-MAY-2002 (first entry)  
 XX DE Canine distemper virus cDNA.  
 XX KW CDV; ss; antiviral; immunostimulant; immunogen; vaccine; gene therapy;  
 XX KW progressive demyelinating disease; Paget's disease; multiple sclerosis.  
 XX OS Canine distemper virus.  
 XX PN WO200200883-A2.  
 XX PD 03-JAN-2002.  
 XX PF 22-JUN-2001; 2001WO-US020157.  
 XX PR 23-JUN-2000; 2000US-0213698P.  
 XX PA (AMCY ) AMERICAN CYANAMID CO.

PI Parks CL, Sidhu MS, Walpita P, Kovacs GR, Udem SA;  
 XX WPI; 2002-130895/17.

XX Producing recombinant canine distemper virus (CDV), useful in gene  
 PT therapy, comprises transfecting cell with transcription vector containing  
 PT nucleic acid encoding CDV genome or antigenome.

XX Claim 43; Fig 6; 109pp; English.

XX The invention relates to producing a recombinant canine distemper virus  
 CC (CDV) comprising transfecting a cell with a composition  
 CC containing: (a) a transcription vector containing an isolated nucleic  
 CC acid molecule encoding a genome/antigenome of CDV, or its variant  
 CC sequence; and (b) an expression vector with an isolated nucleic acid  
 CC molecule encoding the trans-acting proteins (N, P and L) for  
 CC encapsidation, transcription and replication. The recombinant CDV is used  
 CC as an immunogen for immunising an animal or human against CDV. Also  
 CC included are a plasmid comprising a CDV genome/antigenome or encoding CDV  
 CC proteins, and a host cell transformed with the vector. The recombinant  
 CC virus are useful as vectors for expressing foreign genetic information,  
 CC e.g. foreign genes, for applications including immunogenic or  
 CC pharmaceutical compositions for pathogens other than canine distemper,  
 CC gene therapy, and cell targeting. The recombinant CDVs are also used in  
 CC generating antibodies, prophylactic and therapeutic applications,  
 CC specifically in treating or ameliorating canine distemper infection  
 CC (which causes a progressive demyelinating disease with neurological  
 CC symptoms and is implicated in human multiple sclerosis and Paget's  
 CC disease), and in preparing mutant virus and immunogenic compositions.  
 CC Protein and nucleotide sequences may be used to design screening systems  
 CC for compounds that interfere or disrupt normal virus development, via  
 CC encapsidation, replication, or amplification. The present sequence is the  
 CC Canine distemper virus entire cDNA sequence

XX Sequence 15690 BP; 4810 A; 3355 C; 3387 G; 4138 T; 0 U; 0 Other;

Query Match 7.3%; Score 106.6; DB 6; Length 15690;  
 Best Local Similarity 46.0%; Pred. No. 1.1e-21;  
 Matches 361; Conservative 0; Mismatches 424; Indels 0; Gaps 0;  
 Qy 499 TCTATCCAGGCTCAGGTATGGTCCAGTAGCGAAGCGCATGACTGCATATGACAGCA 558  
 Db 612 TCCATCTGGCTCAATTTGGATCCTGCTAGTAAGCGGTGACTGCTCTCTGATCTGCA 671  
 Qy 559 GATGAGTCGGAACAAGAGAAATCAATAGTACATGCGAGGAGGAGTCCAGAGAGAG 618  
 Db 672 GCCGACTCGGAGATGAGAAAGGTGGAATTAAGTATACCCAGCAAGACGCTGGTCGAGAA 731  
 Qy 619 TACATCTCCACCTGTATGCGAGGTGCAATTCAACTCAACATCAGACATTTCTCTGGCA 678  
 Db 732 TTTAGATGAACAAATCTGGCTTGATATTGTTAGAAACAGGATTCGTGAGGACCTATCT 791  
 Qy 679 GTCCGCATTTTCTTAGTTAGCGAGCTTAAGAGAGCGCGCAATACGGCAGGTGGAGCTCC 738  
 Db 792 TTGAGGGGATTCATGGTGGCGCTCATCTTGACATCAAAACGATCCCGAGGAAACAAAGCCT 851  
 Qy 739 ACGTATTACAACTTAGTAGGGGATGTAGATCATACATCAGGAACACCGGACTTACTGCA 798  
 Db 852 AGAATTCGTGAAATGATTTGTGATATAGATAACTACATTTGTGGAAGCTGGGTAGCTAGT 911  
 Qy 799 TTCTTCTTACACTCAAAATATGGAATTAATACCAAGACATCAGCCCTAGCACTCAGCAGC 858  
 Db 912 TTCATCTTAATCAAGTTTGGCATTTGAACTATGATATCGGCTCTTGGGTTGCATGAG 971  
 Qy 859 CTCACAGCGGATATCCAAAAGATGAAGCAGCTCATGCGTTTATATCGGATGAAGGAGAA 918  
 Db 972 TTTTCGGAGAAATTAACAACATTTGAACTCCCTCATGATGCTATATCAACAGATGGGTGAA 1031  
 Qy 919 AATGCGCGTACATGACATTTGCTAGGTGACAGTGCATCAGATGAGCTTTGACACCGCTGAG 978  
 Db 1032 ACAGCACCGTACATGTTATCTTGGAAACTCTGTTCAAAACAATTTAGTGCAGGCTCC 1091  
 Qy 979 TATGCACAGCTTTATCTTTTGGCATGGGATGGCATCAGTCTTAGATAAAGGAACCTGGC 1038

```

Db 1092 TACCCATTGCTCTGGAGTTATGCTATGGGGTTGGTGTCAAACTTGAAGAACTCCATGGGA 1151
QY 1039 AATATCAATTCGCCAGAGACTTCATGACACATCTTCTGGAGACTCGGGTGGAGTAT 1098
Db 1152 GGGTTAAATTCGGTCGATCTTACTTTGACCCAGCTTACTTCAGACTCGGGCAAGAAATG 1211
QY 1099 GCTCAGGCTCAGGGAGTAGCATCAACGAAGACATGCTGCTGAGCTAAACTAAACCCCG 1158
Db 1212 GTTAGGAGATCTGCCGCAAGTAAGCTCTGCATTTGCCGCGAGCTTGGCATCACAAG 1271
QY 1159 GCAGCAAGAGGGGCTCGGAGCTGCTGCTGCCAACAGAGTGTCTGAGGAAACTGGCAGCGTG 1218
Db 1272 GAGGAGCTCAGCTAGTGTGAGAAATAGCATCCAGACAAACAGAGACCGGACAATTCGA 1331
QY 1219 GATATTCTTACTCAACAGCGGGTCTCTCACTGGGCTCAGCATGAGGCCCCCGAGCC 1278
Db 1332 GCTACTGGTCTTAAGCAATCCCAATCACTTTTCTGCACTCGGAAAGATCCGAAATCGCC 1391
QY 1279 TCTCA 1283
Db 1392 AATCA 1396

```

## RESULT 14

ABK15000  
ID ABK15000 standard; DNA; 15690 BP.

AC ABK15000;

XX 08-MAY-2002 (first entry)

XX Canine distemper virus positive, antigenomic strand.

XX CDV; ss; antiviral; immunostimulant; antigenomic strand; immunogen;

XX vaccine; gene therapy; progressive demyelinating disease;

XX Page's disease; multiple sclerosis.

XX Canine distemper virus.

XX WO200200883-A2.

XX 03-JAN-2002.

XX 22-JUN-2001; 2001WO-US020157.

XX 23-JUN-2000; 2000US-0213698P.

XX (AMCY ) AMERICAN CYANAMID CO.

XX Parks CL, Sidhu MS, Walpita P, Kovacs GR, Udem SA;

XX WPI; 2002-130895/17.

XX Producing recombinant canine distemper virus (CDV), useful in gene

XX therapy, comprises transfecting cell with transcription vector containing

XX nucleic acid encoding CDV genome or antigenome.

XX Claim 21; Page; 109pp; English.

XX The invention relates to producing a recombinant canine distemper virus

XX (CDV) comprising transforming/transfecting a cell with a composition

XX containing: (a) a transcription vector containing an isolated nucleic

XX acid molecule encoding a genome/antigenome of CDV, or its variant

XX sequence; and (b) an expression vector with an isolated nucleic acid

XX molecule encoding the trans-acting proteins (N, P and L) for

XX encapsidation, transcription and replication. The recombinant CDV is used

XX as an immunogen for immunising an animal or human against CDV. Also

XX included are a plasmid comprising a CDV genome/antigenome or encoding CDV

XX proteins, and a host cell transformed with the vector. The recombinant

XX virus are useful as vectors for expressing foreign genetic information,

XX e.g. foreign genes, for applications including immunogenic or

XX pharmaceutical compositions for pathogens other than canine distemper,

CC gene therapy, and cell targeting. The recombinant CDVs are also used in

CC generating antibodies, prophylactic and therapeutic applications,

CC specifically in treating or ameliorating canine distemper infection,

CC (which causes a progressive demyelinating disease with neurological

CC symptoms and is implicated in human multiple sclerosis and Paget's

CC disease), and in preparing mutant virus and immunogenic compositions.

CC Protein and nucleotide sequences may be used to design screening systems

CC for compounds that interfere or disrupt normal virus development, via

CC encapsidation, replication, or amplification. The present sequence is the

CC canine distemper virus positive, antigenomic strand. Note: The present

CC sequence is not shown in the specification but was obtained from Genbak

CC using accession number AF014953

XX

XX SQ

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 7.3%; Score 106.6; DB 6; Length 15690;  
Best Local Similarity 46.0%; Pred. No. 1.1e-21;  
Matches 361; Conservative 0; Mismatches 424; Indels 0; Gaps 0;

```

QY 499 TCTATCCAGGCTCAGGTATGGGTACAGTAGCGAAGGCCATGACTGCATATGAGACGCA 558
Db 612 TCACTTGGCTCAAAATTTGGATCTCTAGCTAAGCGGTGACTCTCTGTGATCTGCA 671
QY 559 GATGAGTCGGAACCAAGAGAAATCAATAGTACATGTCAGCAAGCGCAGAGTCCAGAGAAG 618
Db 672 GCCGACTCGGAGATGAGAAAGTGGATTAGTATATCCAGCAAAAGAGCTGTGGTCGGAGAA 731
QY 619 TACATCCCTCCACCTGTATGAGGAGTGCATTAATCACTCAATCAGATCAATCTCTGCGCA 678
Db 732 TTTAGAAATGAACAAATCTGGCTTGATATTTAGAAACAGGATTTGCTGAGGACCTATCT 791
QY 679 GTCCGCAATTTCTTAGTTAGCGAGCTTAAAGAGAGCGCGCAATACGCGAGTGGAGCTCC 738
Db 792 TTGAGCGGATTCATGCTGGCGCTCATCTTGGACATCAACAGATCCCGAGGAACAGCCT 851
QY 739 AGTATTACAACTTAGTAGGGGATGAGACTCATATCAGCAAGAACCCGAGCTTACTGCA 798
Db 852 AGAATTGCTGAATGATTTGTGATAGATACTACATTTGTGGAAGCTGGTGGTAGTAGT 911
QY 799 TTCTTCTTACACTCAATATGGATTAATCAACAGACATCAGCCCTAGCACTCAGCAGC 858
Db 912 TTCACTCTAACTATCAAGTTTGGCAATGAAACTATGATATCCGGCTCTTGGGTGTCATGAG 971
QY 859 CTCACAGCGATATCCAAAAGATGAAGCAGCTCATGCGTTTATATCGGATGAAGGAGAA 918
Db 972 TTTTCCGGAGATTAACAACTATTGATCCCTCATGATGCTATATCAACAGATGGGTGAA 1031
QY 919 AATGCGCGTACATGACATTTGCTAGGTGACAGTGAATGAGCTTTGACCCGCTGAG 978
Db 1032 ACAGCAGCGTACATGTTATCTTTGAAAACCTCTGTTCAAAACAAATTTAGTCAGGGTCC 1091
QY 979 TATGACAGCTTTATTCTTTTGGCATGGCATGGCATCAGTCTTAGATAAAGAACTGGC 1038
Db 1092 TACCATGCTCTGGAGTTATGCTATGGGGTGTGTTGAACTTTGAAACTCCATGGGA 1151
QY 1039 AAATACCAATTCGCCAGAGACTTTCATGAGCACTCATTTCTGGAGACTCGGGGTGGAGTAT 1098
Db 1152 GGGTTAAATTTGGTCTGATCTTACTTTGACCCAGCTTACTTCAGATCGGGCAAGAAATG 1211
QY 1099 GCTCAGGCTCAGGGAGTAGCATCAACGAAGACATGGCTGTGAGCTAAAATAACCCCG 1158
Db 1212 GTTAGGAGATCTGCCGCAAGTAAGCTCTGACCTTGGCGCGAGCTTGGCATCACCAG 1271
QY 1159 GCAGCAAGAGGGGCTCGCAGCTGCTGCCCAACAGTGTCTGAGGAAACTGGCAGCGTG 1218
Db 1272 GAGGAAGCTCAGTAGTGTGAGAAATAGCATCCAAAGACAAACAGAGGACCGGACAATCGA 1331
QY 1219 GATATTCTTACTCAACAGCGGGGTCTCTCACTGGGCTCAGCGATGAGGCCCCCGAGCC 1278
Db 1332 GCTACTGCTCTAAGCAATCCCAATCACTTTTCTGCACTCGGAAAGATCCGAAAGTCGCC 1391
QY 1279 TCTCA 1283

```



Db 1392 AATCA 1396

RESULT 15

ABK15039

ABK15039  
ID ABK15039 standard; DNA: 18826 BP.

AC ABK15039:

08-MAY-2002 (first entry)

XX DE Canine distemper virus full length genomic clone plus vector.

CDV; ds; antiviral; immunostimulant; immunogen; vaccine; gene therapy;  
KW KW progressive demyelinating disease; Pacet's disease; multiple sclerosis;  
XX

XX OS Canine distemper virus.

XX	Key	Location/Qualifiers
FH		

FH	key	Location/
FT	misc feature	1..2198

```
FT      msec_reacure      1. .2130
FT      /*taq= a
```

```

FT /cay= a
FT /note= "Vector sequence"

```

FI	misc feature	178889
FT		178889

```
FI misc_feature 1/803: .I
FT /*taq= b
```

```
FTT /note= "Vector sequence"
```

AX  
PN  
WO200200883-A2.

03-JAN-2002.

22-JUN-2001: 2001WO-US020157-XX PF

23-JUN-2000 2000US-0213698P

XX  
PA (MCCY ) AMERICAN CYANAMID CO

XX	DT	Darks AT	Slabbs MC	Walnuts D	Youngs CD	Idem SA:
----	----	----------	-----------	-----------	-----------	----------

XX 12/12

XX Producing recombinant canine distemper virus (CDV), useful in gene  
PT therapy, comprises transfecting cell with transcription vector containing  
PT nucleic acid encoding CDV genome or antigenome.

PS Claim 43: Fig 7: 109pp: English.

The invention relates to producing a recombinant canine distemper virus (CDV) comprising transforming/transfecting a cell with a composition containing: (a) a transcription vector containing an isolated nucleic acid molecule encoding a genome/antigenome of CDV, or its variant sequence; and (b) an expression vector with an isolated nucleic acid molecule encoding the trans-acting proteins (N, P and L) for encapsidation, transcription and replication. The recombinant CDV is used as an immunogen for immunising an animal or human against CDV. Also included are a plasmid comprising a CDV genome/antigenome or encoding CDV proteins, and a host cell transformed with the vector. The recombinant virus are useful as vectors for expressing foreign genetic information, e.g. foreign genes, for applications including immunogenic or pharmaceutical compositions for pathogens other than canine distemper, gene therapy, and cell targeting. The recombinant CDVs are also used in generating antibodies, prophylactic and therapeutic applications, specifically in treating or ameliorating canine distemper infection (which causes a progressive demyelinating disease with neurological symptoms and is implicated in human multiple sclerosis and Paget's disease), and in preparing mutant virus and immunogenic compositions. Protein and nucleotide sequences may be used to design screening systems for compounds that interfere or disrupt normal virus development, via encapsidation, replication, or amplification. The present sequence is the Canine distemper virus entire genomic DNA sequence plus vector sequences.

XX  
SQ , sequence 18826 BP: 5565 A: 4157 C: 4188 G: 4916 T: 0 U: 0 Other:

### Query Match

7.34: Score 106.6: DB 6: Length 18826:

Best Local Similarity 46.0%; Pred. No. 1.2e-21;  
Matches 361; Conservative 0; Mismatches 424; Indels 0; Gaps 0;

Qy	499	TCTATCCAGGCTCAGGTATGGGTACAGTAGCGAAGGCCATGACTGTCATATGAGACAGCA	558
Db	2810	TCCATCTTGGCTCAAAATTTTGGATCTCTGCTAGCTTAAAGCGGTGACTGCTCTGTACTGCA	2869
Qy	559	GATGAGTCGGAAACAAGAAATCAATAAGTACATGCAAGCAGGCAAGTCCAGAAAG	618
Db	2870	GCCGACTCGGAGATGAGAAGTGGATTAACTATACCCAGCAAGACGTGTGTCGGAGAA	2929
Qy	619	TACATCTCTCACCCCTGTATGCGAGGTGCAATTCAACTCAATCAGACATTTCTCTGGCA	678
Db	2930	TTTAGAATGAACAAAATCTGGCTTGATATTGTTAGAAACAGGATTGCTGAGGACCTATCT	2989
Qy	679	GTCCGCATTTTCTTAGTTAGCTAGGAGCTTAAAGAGGGCGCAATACGGCAGGTGGGAGCTCC	738
Db	2990	TTGAGGCGGATTCATGTGTGGCGCTCATCTTGGACATCAACGATCCCGAGGAACAAGCCT	3049
Qy	739	ACGTATTACAACTTTAGTATGGGGATGTAGACTCATACATCAGGAACACCGGACTTTACTGCA	798
Db	3050	AGAAATGCTGAATGATTTGTGATATAGATAAACTACATTGTGGAAGCTGGTTAGCTAGT	3109
Qy	799	TTCTTCTCTTACACTCAAAATATGGAATTAATACCAAGACATCAGCCCTAGCACTCAGCAGC	858
Db	3110	TTCATCTTAATCATCAAGTTTGGCAATTGAAACTATGTATCCGGCTCTTGGTGTGCATGAG	3169
Qy	859	CTCACAGCGCATATCCAAAAGATGAAGCAGCTCATGCGTTTATTCGATGAAGGAGGAGAA	918
Db	3170	TTTTTCCGGAGAAATTAACAACTATTGAAATCTCCTCATGATGCTATATCAACAGATGGGTGA	3229
Qy	919	AATGCGCGCTACATGACATTTCTAGGTGACAGTGATCAGATGAGCTTTGCACCGGTGAG	978
Db	3230	ACAGCACGTTACATGGTTATCTTTGGAAAACTCTGTTCAAAACAAATTTAGTCAGCGGTCC	3289
Qy	979	TATGCAAGCTTTATCTTTTGGCATGGGCATGGCATCAGTCTTAGATTAAGGAACCTGGC	1038
Db	3290	TACCATTGTCTGGAGTTATGCTATGCGGGTGTGTTGAACTTGAAAACTCCATGGGA	3349
Qy	1039	AAATACCAATTTGCGCCAGAGACTTCATCAGCACATCATTTCTTGGAGACTCGGGGTGGAGTAT	1098
Db	3350	GGGTAAATTTGCTGCACTTACTTTTGAACCCAGCTTACTTCAGCTCTGGGCAAGAAATG	3409
Qy	1099	GCTCAGGCTCAGGGAGTAGCATCAACGAAGACATGGCTGCTAGCTTAAACTTAACCCCG	1158
Db	3410	GTTAGGAGATCTGCGCGCAAGTAAGCTCTGCACTTTCGCGCGAGCTTGGCATCACCAAG	3469
Qy	1159	GCAGCAAGAGGGGCTTGGCAGCTGCTGCCAACGAGTGTCTGAGGAAACTGGCAGCGTG	1218
Db	3470	GAGGAAGCTCAGCTTAGTGTGCAGAAATAGCATCCAAAGCAACAGAGGACCGGACAAATTCGA	3529
Qy	1219	GATATTCTTACTCAACAAGCGGGTCTCCTCACTGGGCTCAGCGATGAGGCGCCCGAGCC	1278
Db	3530	GCTACTGGTCTTAGCAATCCCAATCACTTTTCTGCACTCGGAAAGATCCGAAGTCCGC	3589
Qy	1279	TCTCA 1283	
Db	3590	AATCA 3594	

Search completed: September 3, 2004, 20:13:21  
Job time : 867 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 18:10:20 ; Search time 9149 Seconds  
(without alignments)  
6964.068 Million cell updates/sec

Title: US-09-970-851-1  
Perfect score: 1470  
Sequence: 1 atgtcttcgattcatgatg.....acaccagctgggggatttga 1470

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hcg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sv.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vl.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pre: No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1470	100.0	1470	14	AF284646 Newcastle
2	1302	88.6	1746	14	AF419402 Newcastle
3	1286	87.5	1700	14	AF419404 Newcastle
4	1271.6	86.5	1746	14	AF419405 Newcastle
5	1270	86.4	1580	14	NDI306303 Newcastle
6	1268.4	86.3	1474	14	AF419395 Newcastle
7	1268.4	86.3	1557	14	AF419406 Newcastle
8	1265.2	86.1	1580	14	NDI306302 Newcastle
9	1257.2	85.5	1488	14	AF419407 Newcastle
10	1257.2	85.5	1585	14	AF419396 Newcastle
11	1257.2	85.5	1694	14	AF419401 Newcastle
12	1252.4	85.2	1561	14	AF419410 Newcastle
13	1250.8	85.1	15192	14	AF431744 Newcastle
14	1250.8	85.1	15192	14	AF473851 Newcastle
15	1246	84.8	1694	14	AF419400 Newcastle
16	1242.8	84.5	1575	14	AF419398 Newcastle
17	1242.8	84.5	1694	14	AF419399 Newcastle
18	1230	83.7	1746	14	NDVNCPR Newcastle
19	1220.4	83.0	2617	14	PANDV3E Newcastle
20	1213.6	82.6	1597	14	AF419397 Newcastle
21	1210.8	82.4	1746	14	AF419408 Newcastle
22	1166	79.3	1820	14	AF144730 Newcastle
23	1164.4	79.2	1746	14	AF419409 Newcastle
24	1161.2	79.0	1747	14	AF064091 Newcastle
25	1159.8	78.9	1470	14	AF419403 Newcastle
26	1159.6	78.9	1684	14	NDI306301 Newcastle
27	1159.6	78.9	15186	14	AF375823 Newcastle
28	1158	78.8	1535	14	AF060483 Newcastle
29	1158	78.8	15186	14	AF309418 Newcastle
30	1156.4	78.7	15186	14	NDVY18898 Newcastle
31	1150.2	78.2	1470	14	AF419411 Newcastle
32	1148.4	78.1	15186	6	AX008510 Sequence
33	1148.4	78.1	15186	6	BD218398 Newcastle
34	1148.4	78.1	15186	14	AF077761 Newcastle
35	1146.8	78.0	15186	14	AY225110 Newcastle
36	362.6	24.7	479	14	AF213474S2 Newcastle
37	356.6	22.8	427	14	AF213474S1 Newcastle
38	315.4	21.5	589	14	AF512528 Newcastle
39	304.2	20.7	589	14	AF512530 Newcastle
40	302.6	20.6	589	14	AF512532 Newcastle
41	301	20.5	589	14	AF512529 Newcastle
42	301	20.5	589	14	AF512531 Newcastle
43	299.4	20.4	589	14	AF512533 Newcastle
44	299.4	20.4	589	14	AF512535 Newcastle
45	294.6	20.0	589	14	AF512534 Newcastle

ALIGNMENTS

RESULT 1  
AF284646  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE

AF284646 Newcastle disease virus strain AF2240 nucleocapsid protein (NP)  
gene, complete cds.  
AF284646.1 GI:12382261  
Newcastle disease virus  
Viruses; ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Paramyxovirinae; Rubulavirus.  
1 (bases 1 to 1470)  
Kho,C.I., Tan,W.S. and Yusoff,K.  
Direct Submission

## JOURNAL

Submitted (03-JUL-2000) Biochemistry and Microbiology, University  
Putra Malaysia, Faculty of Science and Environmental Studies,  
Serdang, Selangor 43400, Malaysia

## FEATURES

## source

1..1470  
/organism="Newcastle disease virus"  
/mol\_type="genomic RNA"  
/strain="AF2240"  
/db\_xref="taxon:11176"

## gene

1..1470  
/gene="NP"

## CDS

1..1470  
/gene="NP"

/codon\_start=1  
/product="nucleocapsid protein"

/protein\_id="AG53089.1"

/db\_xref="GI:12382262"

/translation="MSSVDFEYQLAAQTRPNHAGGGERSTLREVPVFTLSDD  
PEDRNFAVFCRIAYSDANKPLRQALISLCSHVSQVRNHEVALAGKNEATITVL  
EIDGFTSSVPNNRSGVSEERQRMWLAGSLPRACNSNGTFFVAGVEDDAPEDITVL  
TLERILSIOQVWVTVAKMTAYETADESETRINKYMQQGRVOKYILHPVCRSAIQ  
LIRHSLAVRIPLVSLKTRNTAGSSFTYVLVDVDSYIENGLTAFPLIKYGIN  
TKTSALASLSDIGDIKKQMLRMLYRKNGENAPYMLLGDSDMSFAPABYAOIYSFA  
MGWASVLDKGYQIPARDFMSTSFWRIGVEYAOAGSSINEDMAELKLTAPARRGL  
AAAAQRVSESGVDIPTQOAGVLTGLSDGPRASQGSNKSQGPDPADGETQFLDL  
MRVANSMEEAPNSAQSTTHPEPPPTPGPSQNDTDWGY"

## ORIGIN

Query Match 100.0%; Score 1470; DB 14; Length 1470;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTCTTCGGTATTCGATGAATACAGCAGCTCCTCGCTGCTCAGACTCGCCCCCAATGGA 60  
DB 1 ATGCTCTTCGGTATTCGATGAATACAGCAGCTCCTCGCTGCTCAGACTCGCCCCCAATGGA 60

QY 61 GCTCAGGAGGGGAGAGAGAGGAGCACTTTAAGAGTTGAGGTCCAGTATTCATCTTT 120  
DB 61 GCTCAGGAGGGGAGAGAGAGGAGCACTTTAAGAGTTGAGGTCCAGTATTCATCTTT 120

QY 121 AACAGTGCAGTCCAGAAATAGTGGATTTTTCGGTATTCCTGCTTCGGATTCGTT 180  
DB 121 AACAGTGCAGTCCAGAAATAGTGGATTTTTCGGTATTCCTGCTTCGGATTCGTT 180

QY 181 AGCGAGGAGCGCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCGTGTCCCAT 240  
DB 181 AGCGAGGAGCGCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCGTGTCCCAT 240

QY 241 TCTCAAGTGCAGAGAACCATGTTGCCCTTGAGGAAACAGAAATGAGGCTACACTGCT 300  
DB 241 TCTCAAGTGCAGAGAACCATGTTGCCCTTGAGGAAACAGAAATGAGGCTACACTGCT 300

QY 301 GTTCTTGAGTGCAGTGGTTTACAGCAGCGTGCCTCAGTTCAAACAGGAGTGGGTG 360  
DB 301 GTTCTTGAGTGCAGTGGTTTACAGCAGCGTGCCTCAGTTCAAACAGGAGTGGGTG 360

QY 361 TCTGAGGAGAGCAGACAGAGATTCATGTTGATGAGCAGGCTCTCTCCCTCGGCGTGCAGT 420  
DB 361 TCTGAGGAGAGCAGACAGAGATTCATGTTGATGAGCAGGCTCTCTCCCTCGGCGTGCAGT 420

QY 421 AACGGTACTCGTTCGTTCAGCGCTGGGTTGAAGATGATGACACCAAGAGATATCACTGAT 480  
DB 421 AACGGTACTCGTTCGTTCAGCGCTGGGTTGAAGATGATGACACCAAGAGATATCACTGAT 480

QY 481 ACTCTGGAAGAATCCTGCTATCCAGGCTCAGGTATGGTTCACAGTACGAGGCCATG 540  
DB 481 ACTCTGGAAGAATCCTGCTATCCAGGCTCAGGTATGGTTCACAGTACGAGGCCATG 540

QY 541 ACTGCATATGACAGCAGATGAGTCGGAACCAAGAGAAATCAATAAGTACATGAGCAAA 600  
DB 541 ACTGCATATGACAGCAGATGAGTCGGAACCAAGAGAAATCAATAAGTACATGAGCAAA 600

QY 601 GGCAGAGTCCAGAGAGTACATCTCCACCTGTATGAGGAGTGCATTCACATCA 660

DB 601 GGCAGAGTCCAGAGAGTACATCTCCACCTGTATGAGGAGTGCATTCACATCA 660  
QY 661 ATCAGACATTCCTCGCAGTCCGCAATTTCTTAGTTAGCGAGCTTAAGAGAGCCGCAAT 720  
DB 661 ATCAGACATTCCTCGCAGTCCGCAATTTCTTAGTTAGCGAGCTTAAGAGAGCCGCAAT 720  
QY 721 AGGCGAGTGGGAGCTCCAGCTATTACAACTTAGTAGGGGATGTAGACTCATCATCAGG 780  
DB 721 AGGCGAGTGGGAGCTCCAGCTATTACAACTTAGTAGGGGATGTAGACTCATCATCAGG 780  
QY 781 AACACCGGACTTACTGCAATTTCTTACATCTCAATATGGAATTAATACCAAGACATCA 840  
DB 781 AACACCGGACTTACTGCAATTTCTTACATCTCAATATGGAATTAATACCAAGACATCA 840  
QY 841 GGCCTAGCACTCAGCAGACCTCAAGCGGATATCCAAAGATGAAGCAGCTCATGCGTTTA 900  
DB 841 GGCCTAGCACTCAGCAGACCTCAAGCGGATATCCAAAGATGAAGCAGCTCATGCGTTTA 900  
QY 901 TATCGGATCAAGGGAGAAATGCGCGTACATGACATTCCTAGTGCACAGTGCATGATG 960  
DB 901 TATCGGATCAAGGGAGAAATGCGCGTACATGACATTCCTAGTGCACAGTGCATGATG 960  
QY 961 AGCTTTGCAACCGCTGAGTATGCAAGCTTTTATCTTTTCCATGGGCATGGCATCAGTC 1020  
DB 961 AGCTTTGCAACCGCTGAGTATGCAAGCTTTTATCTTTTCCATGGGCATGGCATCAGTC 1020  
QY 1021 TTAGATAAAGAACTGGGAAATACCAATTCGCGCAGAGACTTCATGAGCAGATCTCTGG 1080  
DB 1021 TTAGATAAAGAACTGGGAAATACCAATTCGCGCAGAGACTTCATGAGCAGATCTCTGG 1080  
QY 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTGCATCAACGAGACATGCTGCT 1140  
DB 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTGCATCAACGAGACATGCTGCT 1140  
QY 1141 GAGCTAAAACCTAACCCCGGAGAGAGGGGCTGGCAGTCTGCCCAAACGAGTGTCT 1200  
DB 1141 GAGCTAAAACCTAACCCCGGAGAGAGGGGCTGGCAGTCTGCCCAAACGAGTGTCT 1200  
QY 1201 GAGGAACTGGGAGCGTGGATTTCTACTAAAGCCGGGCTCTCTCTGGGCTCAGC 1260  
DB 1201 GAGGAACTGGGAGCGTGGATTTCTACTAAAGCCGGGCTCTCTCTGGGCTCAGC 1260  
QY 1261 GATGAGGCGCCCGGAGCTCTCAGGCTGATCGAAGTCCGAGGGCAACAGATGCC 1320  
DB 1261 GATGAGGCGCCCGGAGCTCTCAGGCTGATCGAAGTCCGAGGGCAACAGATGCC 1320  
QY 1321 GGAGATGGGGAGACCAATTTCTTGGATTTGATGAGCAGTGGCGAAACAGCATGCGAGAA 1380  
DB 1321 GGAGATGGGGAGACCAATTTCTTGGATTTGATGAGCAGTGGCGAAACAGCATGCGAGAA 1380  
QY 1381 GGGCCAAACTCGGACAGAGCAACCCACCGGAAACCCCGGAGTCCCGGGCCATCA 1440  
DB 1381 GGGCCAAACTCGGACAGAGCAACCCACCGGAAACCCCGGAGTCCCGGGCCATCA 1440  
QY 1441 CAAGATAACGACACCGACTGGGGTATTGA 1470  
DB 1441 CAAGATAACGACACCGACTGGGGTATTGA 1470

## RESULT 2

AF419402

## LOCUS

DEFINITION

Newcastle disease virus chicken/Italy/Milano/45 nucleocapsid

protein gene, complete cds.

ACCESSION

AF419402

VERSION

AF419402.1

KEYWORDS

ORGANISM

REFERENCE

1746 bp RNA linear VRL 21-PEB-2002  
Newcastle disease virus  
Newcastle disease virus  
Newcastle disease virus  
Viruses; ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Paramyxovirinae; Rubulavirus.  
1 (bases 1 to 1746)



VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AF419404.1 GI:15811675  
Newcastle disease virus  
Newcastle disease virus  
Viruses; ssRNA negative-strand viruses; Mononegavirales;  
Paramyxoviridae; Paramyxovirinae; Rubulavirus.

REFERENCE  
AUTHORS  
TITLE

Seal, B.S., Crawford, J.M., Sellers, H.S., Locke, D.P. and King, D.J.  
(bases 1 to 1700)  
Nucleotide sequence analysis of the Newcastle disease virus  
nucleocapsid protein gene and phylogenetic relationships among the  
Paramyxoviridae

JOURNAL  
MEDLINE  
PUBMED

Virus Res. 83 (1-2), 119-129 (2002)

REFERENCE  
AUTHORS  
TITLE

Seal, B.S., Crawford, J.M., Sellers, H.S., Locke, D.P. and King, D.J.  
Direct Submission  
Submitted (14-SEP-2001) Agricultural Research Service, USDA,  
Southeast Poultry Research Laboratory, 934 College Station Road,  
Athens, GA 30605, USA

FEATURES  
Source

1..1700  
/organism="Newcastle disease virus"  
/mol\_type="genomic RNA"  
/isolate="mixed species/U.S.(FL)/Largo/71"  
/db\_xref="taxon:11176"  
67..1536  
/codon\_start=1  
/product="nucleocapsid protein"  
/protein\_id="AAL09057.1"  
/db\_xref="GI:15811675"  
/translation="MSSVFDEYEOILAAOTRPNAGHGGEKSGTLKVPVPTILNSDD  
PEDRWNSVFCRLJAVSEDAKPLRQGLISLCSHSONMHNVALACKNEATLAIL  
RIDGETNSQPNNRSGVSEARQPFVWIASLRACNSGTFPVTAGVEDDAPREDIT  
TLERILSQQVWTVAKAMTAYETADESETRINKYMQQKRVQKYLILHPVCRSAIQ  
LTIHLSLAVRIFLPSVSELKRGNTAGSSTYVNLGVDSYIRNTGLTFAFFLTKYGIN  
TKTSALSLSTGDIQKMKQLMRLYRMKGENAPYMLLIGSDSDMSFAPRYAQLYSFA  
MMAASVLDKQKQKQFARDPMSFTFWRLGVEYAOAHGIIIEDLAADIILTPAARRRL  
TAAQVSEEEIGSDVILLIQQAVFTFLSDGGPOAPQGGNLNRVQGRPEAGVGETFLDL  
MRVANSNREVNQNSAPSTHSELPTTPGSPDNDTWGY"

CDS

ORIGIN

Query Match 87.5%; Score 1286; DB 14; Length 1700;  
Best Local Similarity 92.2%; Pred. No. 0;  
Matches 1355; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 1 ATGTCTTCCTGATTCGATGAATACGAGCAGCTCTCGCTGCTCAGACTCGCCCCAATGGA 60  
DB |||||  
QY 67 ATGTATCCGATTTTGCAGTAATACGAGCAACTCTCTCGCTCAGACTCGCCCCAATGGA 126  
DB |||||  
QY 61 GCTCACGGAGGGGAGAGAGAGGAGGAGCACTTTAAGAGTTGAGTCCCAAGTATTCACCTCT 120  
DB |||||  
QY 127 GCTCACGGAGAGAGAGAGAGGAGGAGCACTTTAAGTTGAGTCCCAAGTATTCACCTCT 186  
DB |||||  
QY 121 AACAGTGACGATCCAGAGAGATAGATGGAAATTTTCGGTATTCTGTCTTCGAGATTCTGTT 180  
DB |||||  
QY 187 AATAGTGACGATCCAGAGATAGATGGAAATTTTCGGTATTCTGTCTTCGGATTGCCCTT 246  
DB |||||  
QY 181 AGCGAGGACCCCAACACCCCTCAGGCAAGTGTCTCTATATCCCTCTGTCTCCCAT 240  
DB |||||  
QY 247 AGCGAGGATGCCAACAAGCCACTCAGGCAAGTGTCTCTATATCCCTCTGTCTCCCAT 306  
DB |||||  
QY 241 TCTCAAGTGATGAGGAAACCATGTTCCTCTGAGGAAACAGAAATGAGGCTACACTGACT 300  
DB |||||  
QY 307 TCTCAAGTGATGAGGAAACCATGTTCCTCTGAGGAAACAGAAATGAGGCTACACTGACT 366  
DB |||||  
QY 301 GTTCTTGAGATCGATGGTTTACAGAGAGCGTGCCTCAGTTCAACAAACAGGAGTGGGTG 360  
DB |||||  
QY 367 ATTCTTGAGATTGATGGTTTACCAACAGCAGCGCCCAAGTTTCAACAAACAGGAGTGGAGTG 426  
DB |||||  
QY 361 TCTGAGGAGAGGACACAGAGATTCATGTGTGATGACAGGATCTCTCCCTCGGCGGTGAGT 420  
DB |||||  
QY 427 TCCGAGGAGAGGACACAGAGATTCATGTGTGATGACGCGGTCTCTCCCTCGGCGGTGAGT 486  
DB |||||

QY 421 AACGGTACTCCGTTGCTCAOCCGCTGGGGTGAAGATGATGCACACAGAGATATCACTGAT 480  
DB |||||  
QY 487 AACGGTACTCCGTTGCTCAOCCGCTGGGGTGAAGATGATGCACACAGAGATATCACTGAT 546  
DB |||||  
QY 481 ACTCTGGAAGAAATCCTCTGCTATCCAGGCTCAGGATGGGTTCACAGTACGAGAGCCATG 540  
DB |||||  
QY 547 ACTCTGGAAGAAATCCTCTGCTATCCAGGCTCAGGATGGGTTCACAGTACGAGAGCCATG 506  
DB |||||  
QY 541 ACTGATATGACAGACAGAGATGAGTCGGAACAAGAGAAATCAATAAGTACATGAGCAAA 600  
DB |||||  
QY 607 ACTGATATGACAGACAGAGATGAGTCGGAACAAGAGAAATCAATAAGTATATGACGAA 666  
DB |||||  
QY 601 GSCAGAGTCCAGAGAGAGTACATCCTCCACCTGTATCGAGAGTGCATTAATCAACTCACA 660  
DB |||||  
QY 667 GSCAGAGTCCAGAGAGAGTACATCCTCCACCTGTATCGAGAGTGCATTAATCAACTCACA 726  
DB |||||  
QY 661 ATCAGACATTTCTCTGGCAGTCCGCAATTTTCTTAGTTAGCGAGCTTAAAGAGAGCCGCAAT 720  
DB |||||  
QY 727 ATCAGACATTTCTTTGGCAGTTGCAATTTTCTTAGTTAGCGAGCTTAAAGAGAGCCGCAAT 786  
DB |||||  
QY 721 ACGGCAGTGGGAGCTCCAGTATTTTACAATTTAGTTAGGGGATGTAGATCATACATCAGG 780  
DB |||||  
QY 787 ACGGCAGTGGGAGCTCCAGTATTTTACAATTTAGTTAGGGGATGTAGATCATACATCAGG 846  
DB |||||  
QY 781 AACACCGGACTTACTGCAATTTCTCTTACACTCAAAATATGAAATTAATACCAAGACATCA 840  
DB |||||  
QY 847 AACACCGGCTTACTGCAATTTCTTTGACACTCAAAATATGAAATTAATACCAAGACATCG 906  
DB |||||  
QY 841 GCCCTAGCACTCAGCAGCTCCACAGCGGATATCCAAAAGATGAAGCAGCTCATGCTGTTTA 900  
DB |||||  
QY 907 GCTCTTGCACTCAGCAGCTCCACAGCGGATATCCAAAATGAAGCAGCTCATGCTGTTTA 966  
DB |||||  
QY 901 TATCGGATGAAGGAGAGAAATGCGCGGTACATGACATTTGCTAGGTGACAGTGTACAGATG 960  
DB |||||  
QY 967 TATCGGATGAAGGAGAGAAATGCAACATACATGACATTTGCTAGGTGACAGTGTACAGATG 1026  
DB |||||  
QY 961 AGCTTTGACCGCTGAGTATGTCACAGCTTTTATCTTTGCTGCGATGGGATGSCATCAGTC 1020  
DB |||||  
QY 1027 AGCTTTGACCGCTGAGTATGTCACAGCTTTTATCTTTGCTGCGATGGGATGSCATCAGTC 1086  
DB |||||  
QY 1021 TTAGATAAAGGAATGCGCAATACCAATTCGCCAGAGACTTTCAGGACACATCATTTCTGG 1080  
DB |||||  
QY 1087 TTAGATAAAGGAATGCGCAATACCAATTCGCCAGAGACTTTCAGGACACATCATTTCTGG 1146  
DB |||||  
QY 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGATGACATCAACAGAGACATGCGTCTGCT 1140  
DB |||||  
QY 1147 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGATTTACCATCATTTGAAGAGCTTGGCTGCT 1206  
DB |||||  
QY 1141 GAGCTTAAACTAAACCCCGGAGCAAGAGGGGCTGGCAGCTGCTGCCCAAGCAGTGTCT 1200  
DB |||||  
QY 1207 GATCTATATCTTAAACCCCGGAGCAAGAGAGGCTGACAGTGTGCTGCCCAAGCAGTGTCT 1266  
DB |||||  
QY 1201 CAGGAAATCGGAGCGTGGATATTCCTACTCAACAGCCGGGCTCTCACTGGGCTCAGC 1260  
DB |||||  
QY 1267 CAGGAAATCGGAGCGTGGATATTCCTATTCAACAGCCGGGCTCTCACTGGACTTAGT 1326  
DB |||||  
QY 1261 GATGGAGCCCGGAGCTCTCAGGTTGATGCAACAGTGCAGAGGCAACCATGCTCC 1320  
DB |||||  
QY 1327 GATGGAGCCCGGAGCTCTCAGGTTGATGCAACAGTGCAGAGGCAACCATGCTCC 1386  
DB |||||  
QY 1321 GGAGATGGGAGAGACCAATTTCTGATTTGATGAGCAGTGGCGAAACAGCATGCGAGAA 1380  
DB |||||  
QY 1387 GGGGTGGAGAGAGACCAATTTCTGATTTGATGAGAGCGGTGGCGAATAGCATGCGAGAA 1446  
DB |||||  
QY 1391 GCGCCAAATCTCGACAGAGACCAACCCCGGAAACCCCGGAGTCTCCCGGGGCAATCA 1440  
DB |||||  
QY 1447 GTACAAATCTCGACAGGACCAACCATTTTCAAGAGCTTCCCGCAACTCTCTGGGCAATCC 1506  
DB |||||  
QY 1441 CAAGATACGACACCGACTGGGGGTATTGA 1470  
DB |||||  
QY 1507 CCAGACACGACACCGACTGGGGGTACTGA 1536  
DB |||||

```

RESULT 4
LOCUS AF419405
DEFINITION Newcastle disease virus chicken/Mexico/37821/96 nucleocapsid
ACCESSION AF419405
VERSION AF419405.1 GI:15811677
KEYWORDS Newcastle disease virus
SOURCE Newcastle disease virus
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales;
Paramyxoviridae; Paramyxovirinae; Rubulavirus.
REFERENCE 1 (bases 1 to 1746)
AUTHORS Seal,B.S., Crawford,J.M., Sellers,H.S., Locke,D.P. and King,D.J.
TITLE Nucleotide sequence analysis of the Newcastle disease virus
nucleocapsid protein gene and phylogenetic relationships among the
Paramyxoviridae
JOURNAL Virus Res. 83 (1-2), 119-129 (2002)
MEDLINE 21854562
PUBMED 11864745
REFERENCE 2 (bases 1 to 1746)
AUTHORS Seal,B.S., Crawford,J.M., Sellers,H.S., Locke,D.P. and King,D.J.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-2001) Agricultural Research Service, USDA,
Southeast Poultry Research Laboratory, 934 College Station Road,
Athens, GA 30605, USA
FEATURES
source
Location/Qualifiers
1..1746
/organism="Newcastle disease virus"
/mol_type="genomic RNA"
/isolate="chicken/Mexico/37821/96"
/db_xref="taxon:111176"
67..1536
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AAL09058.1"
/db_xref="GI:15811678"
/translations="MSVDFEYEQLLAAQTRPNAGHGSGKSTLKVFPVFTLNSDD
PEDRNFVFLRAIYSDSDANKPLRGALISLCSHSQVNRHVLAQKNEATLAIL
EIDGFANSTQFNRRSVSEERAFVIAISLPKACNGTFFVAGVEDDAPEDITD
LTERLSIAQVWTVAKAMTAYETADESSTRINKYMQQGRVQKYLHPVCRSAIQ
LTHSLAVFLVLELRKRGNTAGSSSTYNLVGVDSVIRNRTGTFAPFLTKYGIN
TKTSALASLITGDIQRKMKQLRMLYRKMGENAPYMLLIGSDSDMSFAPAYAQIYRFA
MGSALVDKGTGKQFARDPMSTSFWRGLVEYAAQSSINEDMAELKLTTPARRGL
AAAAQRVSEETGSMIDIPQAGVLTLGSDGPRAPQGSNRPQGPDPAGDGTQFLDL
MRVANSMEAPNSAPGTPPEPPPTPGSPDNDTWGY"
CDS
Query Match 86.5%; Score 1271.6; DB 14; Length 1746;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1346; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 1 ATGTCCTCCGATTTCGATGATACGACGAGCTCCCTCGCTGCTCAGACTCGCCCAATGGA 60
DB 67 ATGTCGTCGATTTCGAGTAGACGACACTCCTCGCTGCTCAGACCCGCCCAATGGA 126
QY 61 GCTCACGGAGGGGAGAGAGAGGAGCAGCTTTAAGAGTTGAGGTCCCAGTATTCACCTTT 120
DB 127 GCTCACGGAGGGGAGAGAGGAGGAGCAGCTTTGAAAGTTGAGGTCCCAGTATTCCTTT 186
QY 121 AACAGTAGCAGTCAGAAAGATAGATGGAATTTTCGGGTATTCCTCTCGGATTCGTT 180
DB 187 AATAGTAGCAGTCAGAAAGATAGATGGAATTTTCGGGTATTCCTCTCTCGGATTCGTT 246
QY 181 AGCGAGAGCGCAACAAACGCTCAGGCAAGGTCTCTCATATCCCTCGTGTGCCAT 240
DB 247 AGTGAGGATGCCAACAAAGCCACTCAGGCAAGGTCTCTCATATCCCTCTGTGTGCCAT 306
QY 241 TCTCAAGTAGTAGGAACCAATGTTGCCCTTTCAGGAAACAAAGATAGGCTTACACTGACT 300
DB 307 TCTCAAGTAGTAGGAACCAATGTTGCCCTTTCAGGAAACAAAGATAGGCTTACACTGCT 366

```

```

QY 301 GTTCTTGAGATCGATGGTGTTCACGACGCGTCCCTCAGTTCAACACAGAGTGGGGTG 360
DB 367 ATTCTTGAGATTCATGGGTTTGTAAACAGACACCTCAGTTCAACACAGAGTGGAGTG 426
QY 361 TCTGAGGAGAGACACAGAGATTCATGGTGATGACGAGGCTCTCTCCCTCGGGGCGTGAGT 420
DB 427 TCCGAGGAGAGACACAGAGATTCATGGTGATGAGTGGGTGGCTCCCTCGGGCATGAGC 486
QY 421 AACGGTACTCCGTTTCGTCACGGCTGGGGTTGAAGATGATGCCACAGAAATATCACTGAT 480
DB 487 AACGGTACTCCGTTTCGTCACAGCGGGGTTGAAGATGATGCCACAGAAATATCACTGAC 546
QY 481 ACTCTGAAAGATCCCTGCTCTATCCAGGCTCAGGTATGGGTCAAGTAGGCAAGAGCCATG 540
DB 547 ACTCTGAAAGATCCCTGCTCTATCCAGGCTCAAGTAGGGTCAAGTAGGCAAGAGCCATG 606
QY 541 ACTGCTATGAGACACAGATGAGTCGGAACAAAGAAAGATCAATAGTAGTACATGCGACAA 600
DB 607 ACTGCTATGAGACACAGATGAGTCAGAAACAAAGAAAGATCAATAGTAGTATGCGACAA 666
QY 601 GGACAGTCCAGAAAGATGATCCTCCACCCCTGTATGACGAGTGCATTTCACTCACA 660
DB 667 GGACAGTCCAGAAAGATGATATCCTCCATCCGCTATGACGAGTGCATTTCACTCACA 726
QY 661 ATCAGACATTCCTGCGCAGTCCGCAATTTCTTAGTTAGCGAGCTTAAGAGAGGCGCAAT 720
DB 727 ATCAGGACCTCTTTGGCAGTTCGCAATTTCTTAGTTAGGAGCTTTAAGAGAGGCGCAAT 786
QY 721 ACGGCAGTGGGAGCTCCAGTATTACAACTTAGTAGGGGATGTAGACTCATCATCAGG 780
DB 787 ACGGCAGTGGGAGCTCCACATATTACAACTTAGTAGGGGATGTAGACTCATCATCAGG 846
QY 781 AACACCGACTTACTGCAATTCCTCTTACACTCACTCAATATGGAATTAATACAGACATCA 840
DB 847 AACACTGGGCGTACTGCAATTCCTCTTGACACTCAATATGGAATTAATACCAAAACATCA 906
QY 841 GCCTTAGCAGCTCAGCAGCCTCAGCGGATATCCAAAGATGAAGCAGCTCATCGGTTTA 900
DB 907 GCTCTGCACTCAGCAGCCTCAGCGGATATCCAAAGATGAAGCAGCTCATCGGTTTA 966
QY 901 TATCGATGAAGGAGAGAAAATGCCCGTACATGACATTCGTTAGTGACAGTGCATGATG 960
DB 967 TATCGATGAAGGAGAGAAAATGCCCGTACATGACATTCGTTAGTGACAGTGCATGATG 1026
QY 961 AGCTTTGCAACCGGCTGAGTATGACAGCTTTATTTCTTTGCGCATGGGCATGCGATCAGTC 1020
DB 1027 AGCTTTGCAACCGGCTGAGTATGACAGCTTTATTTCTTTTGGCATGGGCATGCGATCAGTC 1086
QY 1021 TTAGATAAGGAACTGGCAATACCAATTCGCCAGAGCTTCATGACACATCATCTCTGG 1080
DB 1087 TTAGATAAGGAACTGGCAATACCAATTCGCCAGGACTTTATGACACATCATCTTTTGG 1146
QY 1081 AGACTCGGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAAAGACATGGCTGCT 1140
DB 1147 AGACTCGGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAATGAGGATATGGCTGCC 1206
QY 1141 GAGTAAATACTAACCCCGGAGCAAGAGGGGCTGCGAGCTGCTGCCCAAGAGTGTCT 1200
DB 1207 GAGTAAATACTAACCCCGGAGCAAGAGGGGCTGCGAGCTGCTGCCCAAGAGTGTCT 1266
QY 1201 GAGGAACTCGCAGCGTGGATATTCCTACTCAACAGCCGGGGTCTCTCAGTGGGCTCAGC 1260
DB 1267 GAGGAACTCGCAGCGTGGATATTCCTACTCAACAGCCGGGGTCTCTCAGTGGGCTCAGC 1326
QY 1261 GATGAGGCGCCCGAGCTCTCAGGGGTGGATCGAAAGTTCGAAAGGCAACAGATGCC 1320
DB 1327 GATGAGGCGCCCGAGCTCTCAGGGGTGGATCGAAAGTTCGAAAGGCAACAGATGCC 1386
QY 1321 GAGATGGGAGAGCCCAATTTCTTGGATTTGATGAGAGCTGGCGAAACAGATGCGAGAA 1380
DB 1387 GAGATGGGAGAGCCCAATTTCTTGGATTTGATGAGAGCTGGCGAAACAGATGCGAGAA 1446
QY 1381 GCGCAAACTCCGACACAGAGCAACACCCACCCCGGAAACCCCGGACTCCCGGGCCATCA 1440

```



Db 1447 GCACCAATTCCTGACCGGCACTACCCCTCCAGAGCTCCCCCAACACTGGGCCATCC 1506  
 Qy 1441 CAAGTAACGACACGACTGGGGTATTGA 1470  
 Db 1507 CCAGCAACGACACGACTGGGGGTACTGA 1536

RESULT 5  
 NDI306303  
 LOCUS Newcastle disease virus NP gene for nucleoprotein, genomic RNA, strain 99106. 1580 bp RNA linear VRL 15-MAY-2003  
 DEFINITION  
 AJ306303  
 VERSION AJ306303.1 GI:28268510  
 KEYWORDS NP gene; nucleoprotein.  
 SOURCE Newcastle disease virus  
 ORGANISM Newcastle disease virus  
 Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Paramyxovirinae; Rubulavirus.  
 1  
 Barbezange, C. and Jestin, V.  
 AUTHORS Molecular Characterisation of three avian paramyxovirus type 1  
 TITLE isolated from pigeons in France  
 JOURNAL Virus Genes 26 (2), 175-183 (2003)  
 MEDLINE 22687690  
 PUBMED 12803469  
 REFERENCE 2 (bases 1 to 1580)  
 AUTHORS Barbezange, C.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-FEB-2001) Barbezange C., VIPAC, AFSSA Site de Ploufragan, BP53, 22440 Ploufragan cedex, FRANCE  
 FEATURES  
 source  
 1..1580  
 /organism="Newcastle disease virus"  
 /viralion  
 /mol\_type="genomic RNA"  
 /strain="99106"  
 /specific\_host="racing pigeon"  
 /db\_xref="taxon:11176"  
 /country="France"  
 /note="pigeon paramyxovirus type 1"  
 14..1483  
 /gene="NP"  
 14..1483  
 /gene="NP"  
 /codon\_start=1  
 /product="nucleoprotein"  
 /protein\_id="CAC83753.1"  
 /db\_xref="GI:28268511"  
 /translation="MSSVFDEYEOQLAAOTRPNAGHGEGKSTLKVEVPVPTLNSDD  
 PEDRWNFVFCRLIAVSFEDANKPLROGALISLLCSHVSOMENHVALAGKONEATLAVL  
 EIDSTFSPQPNRSGVSEBAQRFVIAESLPRACSNNTPTFTAGVENVDPEDITD  
 TLERILTSQAQVWTVAKAMATETADESETRINKYMQQRVKQKYLHHPVCRSAIQ  
 LTIHSLAVRFLFSELKRGNTAGGSYYNLVGVDSYIRNTGLTAFPLTKLYGIN  
 TKTSALALSSLTGDIQKMLRLYRMKGENAPYMLILGSDQMSFAPRYAQLYSFA  
 MGMAASVLDKGTQKQFARDFMSTFWRLGVEYAOAGSSINEDMAAEKLTTPAARRGL  
 AAAAQVSEETIGNVDIPTQQAAGVLTGLSKGPRAQGGPSRSSQGPDAQDGETQFLDL  
 MRVANSNREAPDSAQSTIHPEPLPTHGFSQNDTDWGY"

gene

CDS

ORIGIN

Query Match 86.4%; Score 1270; DB 14; Length 1580;  
 Best Local Similarity 91.5%; Pred. No. 0;  
 Matches 1345; Conservative 0; Mismatches 125; Indels 0; Gaps 0;  
 Qy 1 ATGTCTTCGATTTCGATGAATACGACGAGCTCCCTCGCTGCTCAGACTCGGCCCAATGGA 60  
 Db 14 ATGTCATCCGCTTTTTCAGAAATACGACGAGCTCCCTCGCTGCTCAGACTCGGCCCAATGGA 73  
 Qy 61 GCTCACGAGGGGAGAGAGAGGGGAGCACTTTAAGAGTTGAGGTCCCGAGTATTCACCTCT 120  
 Db 74 GCTCACGAGGGGAGAGAGAGGGGAGCACTTTAAGAGTTGAGGTCCCGAGTATTCACCTCT 133

Qy 121 AACAGTGCAGATCCAGAAATAGATGGAATTTTGGGTATTCTGTCTTCGAGTTGCTGTT 180  
 Db 134 AACAGTGCAGATCCAGAAATAGATGGAATTTTGGGTATTCTGTCTTCGAGTTGCTGTT 193  
 Qy 181 AGCGAGGACGCCAACAAACCCGCTCAGGCAAGGTGCTCTCATATCCCTCTGTGCTCCCAT 240  
 Db 194 AGTGAGGATGCCAACAAACCCGCTCAGGCAAGGTGCTCTCATATCCCTCTTATGCTCCCAT 253  
 Qy 241 TCTCAAGTGCAGAGAACCATGTTGCTTCCCTTCAGAGAAACAGAAATGAGGCTACACTGAT 300  
 Db 254 TCTCAAGTGCAGAGAACCATGTTGCTTCCCTTCAGGGAACAGAAATGAGGCTACACTGAT 313  
 Qy 301 GTTCTTGAGATCGATGGTTTTTACACGACGCGCTCAGTTCAAACAACAGAGTGGGGTG 360  
 Db 314 GTTCTTGAGATCGACAGTTTTTACCAACAGTGTGCCCGAGTTCAAACAACAGAGCGGAGTG 373  
 Qy 361 TCTGAGGAGAGACACAGAGATTTCATGTGATAGAGGGTCTCTCCCTCGGCGTGCAGT 420  
 Db 374 TCTGAAGAGAGACACAGAGATTTCATGTGATAGCAGAGTCTCTCCCGGGCATGCACT 433  
 Qy 421 AACGCTACTCCGTTGCTCAGCGCTGGGGTTGAAGATGATGCACCAAGAGATATCACTGAT 480  
 Db 434 AACAAATCTCCGTTGCTTACAGCTGGAGTTGAAAATGATGTGCCCAAGAGATATCACTGAC 493  
 Qy 481 ACTCTGGAAGAATCCTGCTCTATCCAGGCTCAGGCTATGGGTACAGTAGCGAAGGCCATG 540  
 Db 494 ACTCTGGAAGAATCTTATCTATCCAGGCTCAGGCTATGGGTACAGTAGCAAGGCCATG 553  
 Qy 541 ACTGCATATGACAGCAGAGATGAGTTCGGAACAAGAAATCAATAAGTACATGACAGCAA 600  
 Db 554 ACTGCATATGACAGCAGAGATGAGTTCGGAACAAGAAATCAACAAGTATATGACAGCAA 613  
 Qy 601 GGCAGAGTCCAGAGAGATCATCTCCACCTCTGTTCGAGAGTGCATTAACACTCACA 660  
 Db 614 GGTAGAGTCCAGAGAGATCATCTCCATCCGCTATCGAGAGTGCATTAACACTCACA 673  
 Qy 661 ATCAGACATCTCTGGCAGTCCGCAATTTTCTTAGTTACGAGCTTAAGAGAGGCCCAAT 720  
 Db 674 ATCAGACATCTCTAGCAGTCCGCAATTTTCTTAGTTAGTGTAGAGTGAAGGGGGCCCAAT 733  
 Qy 721 ACGGAGTGGGAGTCCACGTAATTAACAATTAGTAGGGGATGTAGACTCATACATCAGG 780  
 Db 734 ACGGAGTGGGAGTCCACGTAATTAACAATTAGTTGGGATGTAGACTCATACATAAGG 793  
 Qy 781 AACACCGGACTTACTGCAATCTTCTCTACACTCAAAATATGGAATTAATACCAAGACATCA 840  
 Db 794 AACACCGGCTTACCGCAATCTTCTCTGACACTCAAGTATGGAATTAATACCAAGACATCA 853  
 Qy 841 GCCCTAGCACTCAGCAGCTCACAGGCGGATATCCAAAAGATGAAGCAGCTCATGCGTTTA 900  
 Db 854 GCCCTTGCACCTCAGCAGCTCACAGGCTGATATCCAAAATGMAACAGCTCATGCGTTTA 913  
 Qy 901 TATCGGATGAAGGAGAGAAATGCGCGTACATGACATTTGCTAGGTGACAGTGTACAGATG 960  
 Db 914 TATCGGATGAAGGAGAGAAATGSCCCGTACATGACATTTGCTAGGTGACAGGACACAGATG 973  
 Qy 961 AGCTTTGACCGGCTGAGTATGACAGCTTTTATCTTTTTCATCGGAGTGCATGCACTGC 1020  
 Db 974 AGCTTCGACCGAGCTGATATGCAAACTTACTCTTTTTCATCGGAGTGCATGCACTGC 1033  
 Qy 1021 TTAGATAAGGAACCTGGCAAAATACCAATTCGCCAGAGACTTTCATGAGACATCACTTCG 1080  
 Db 1034 TTAGATAAGGAACCTGGTAAATACCAATTCGCCAGAGACTTTCATGAGTACATCACTTCG 1093  
 Qy 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAAGATGCTGCTGCT 1140  
 Db 1094 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAATGAGGATATGCTGCTGCT 1153  
 Qy 1141 GAGCTTAAACTAAACCCCGCAGCAGAGAGGGGCTTGGCAGTGTGCTGCCCAACAGAGTGTCT 1200  
 Db 1154 GAGCTGAACCTAACTCCAGCAGCAGAGGGGCTTGGCAGTGTGCTGCCCAACAGAGTGTCT 1213  
 Qy 1201 GAGGAAACTGGCAGCGTGGATATTCCTACTCAACAGCCCGGGTCTCTCACTGGGCTCAGC 1260





Db	1084	AGACTTGGGTAGAGTATGCTCAGGCTCAAGGAGTGCATCAACGAGGACATGCTGCC	1143	ORIGIN
QY	1141	GAGCTAAACTAACCCCGGACGAAAGAGGGGCGCTGGGAGTGTGCTGCCCAACGAGTGTCT	1200	Query Match 86.3%; Score 1268.4; DB 14; Length 1557; Best Local Similarity 91.4%; Pred. No. 0; Matches 1344; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
Db	1144	GAGCTGAACCTAACCCCTGACGAGGAGGGGCGCTGGGAGTGTGCTGCCCAACGAGTGTCT	1203	
QY	1201	GAGGAACTGGCAGCGTGGATATCTCTCAACAGCCGGGCTCTCACTGGGCTCAGC	1260	
Db	1204	GAGGAACTGGTAGCATAAATATCTTACCCCAACAGCCGGGCTCTCACCGGCTCAGC	1263	
QY	1261	GATGAGGCGCCCGGAGCTCTCAGGGTGGATCGAAACAGTCCGAAAGGGCAACAGATGCC	1320	
Db	1264	GACGAGAGCTCCGAGTCTCAACAGCAATCAGACAGTCCGAAAGGGCAACAGATGCC	1323	
QY	1321	GGAGATGGGAGACCCCAATCTTGGATTGTAGAGAGTGGCGAAACAGATGGAGAA	1380	
Db	1324	GGGATGGAGAGACCCCAATCCCTGGACTTGTATGAGAGAGTGGCGAAACAGATGGAGAA	1383	
QY	1381	GGCCCAACTCCGACAGAGCACCCACCCCGGAAACCCCGGACTCCCGGGCCATCA	1440	
Db	1384	GCACCAAACTCCGACAGAGTACCCCAATCTGAGCCCCCCCCCACTCTCGGGCCATCA	1443	
QY	1441	CAAGATAACGACACCGACTGGGGTATGCA	1470	
Db	1444	CAAGCAACGATACTGACTGGGATACTGA	1473	
RESULT 7	AF419406	1557 bp RNA linear VRU 21-FEB-2002		
LOCUS	Newcastle disease virus chicken/Mexico/37822/96 nucleocapsid			
DEFINITION	protein gene, complete cds.			
ACCESSION	AF419406			
VERSION	AF419406.1	GI:15811679		
KEYWORDS	Newcastle disease virus			
SOURCE	Newcastle disease virus			
ORGANISM	Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Paramyxovirinae; Rubulavirus.			
REFERENCE	1 (bases 1 to 1557)			
AUTHORS	Seal, B.S., Crawford, J.M., Sellers, H.S., Locke, D.P. and King, D.J.			
TITLE	Nucleotide sequence analysis of the Newcastle disease virus nucleocapsid protein gene and phylogenetic relationships among the Paramyxoviridae			
JOURNAL	Virus Res. 83 (1-2), 119-129 (2002)			
MEDLINE	21854562			
PUBMED	11864745			
REFERENCE	2 (bases 1 to 1557)			
AUTHORS	Seal, B.S., Crawford, J.M., Sellers, H.S., Locke, D.P. and King, D.J.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-SEP-2001) Agricultural Research Service, USDA, Southeast Poultry Research Laboratory, 934 College Station Road, Athens, GA 30605, USA			
FEATURES	Location/Qualifiers			
source	1. .1557 /organism="Newcastle disease virus" /mol_type="Genomic RNA" /isolate="chicken/Mexico/37822/96" /db_xref="taxon:11176" 4. .1473 /codon_start=1 /product="nucleocapsid protein" /protein_id="AAL09059.1" /translation="MSSVFDEYEQLLAQTNRNAGHGGEKSTLKVFPVFTLNSDD PEDRNWFSVFLRIAYSEDAKPLRQCALISLCSHVSQVRNHEVALAGKONEATLAIL EIDGFANFQFNRSVSEEAQREVIAGSLPRACSNPTFTVAGVDDADPEDITD TLERILSIAQIOWTVAKAMTAVETADESETRINKYMOQGRVOKYKILHPVCRSAIQ LITRHSALVRIPLVSLKRGNTAGGSTYINLVGDVSVIRNTGLTAPFLTLKYGIN TKTSALASLIGDIOKMKQLMRLYRKGNENAPYITLLGSDSDMSFAPAEYAGLYSFA MGWASVLIDGTQIFARDPMSTSFWRILGVYEAQAQGSINEDMAELKILTPAARGL AAAAQRVSEISGVDPYTOQAQGLVTLGSDVGPQASQGSKSKPKGQPNAGDGTQFLDL MRVANSMBREAQNSAPGTTTPPEPPTPGFSPDNDTDWGY"			
CDS				

```
QY 1021 TTAGTAAAGCACTGCGAATACCAATTCGCCAGAGCTTCATGAGCACATCATCTCGG 1080
Db 1024 TTAGTAAAGCACTGCGAATACCAATTCGCCAGGACTTTATGAGCACATCATCTCGG 1083
QY 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAAGACATGGCTGCT 1140
Db 1084 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATTAATGAGGATATGGCTGCT 1143
QY 1141 GAGCTAAACCTAACCCCGGCGAGCAAGAGGGGCTCGGAGCTGCTGCGCCACGAGTGTCT 1200
Db 1144 GAGCTAAACCTAACCCCGGCGAGCAAGAGGGGCTCGGAGCTGCGCTCAACGAGTGTCC 1203
QY 1201 GAGGAACCTGCGAGCTGGATATTCCTACTCAACGAAGCGGGGTCTCACTGGGCTCAGC 1260
Db 1204 GAGGAATTTGCGAGCTGGATGTTCTTACTCAACAGCGCGGGTCTCACTGGACTCAGC 1263
QY 1261 GATGAGGCCCCGAGGCTCTCAGGGTGGATCGCAACAGTGGCAAGGGCAACCATGATGCC 1320
Db 1264 GATGAGGCCCCGAGGCTCTCAGGGCAATTCGAGCAAGCCAAAGGCGCAACCGATGCC 1323
QY 1321 GGAGATGGGAGACCCAAATCTTGATTTGATGAGAGCAGTGGCGAAACAGCATGCGAGAA 1380
Db 1324 GGGACGCGAGAGACCCAAATCTTGATTTGATGAGAGCGGTGGCAAAACAGCATGCGAGAA 1383
QY 1381 GCGCCAAACTCCGACAGACCAACCCACCGGNAACCCCGGACTCCCGGGCCATCA 1440
Db 1384 GCACAAATTTGACCGGGGCACTACCCCTCCAGAGCTCCCGGCAACCTGGGGCATCC 1443
QY 1441 CAAGATAACGACCGACTGGGGGTATTGA 1470
Db 1444 CCAGACAATGACCGACTGGGGGTACTGA 1473
```

```
RESULT 8
NDI306302
LOCUS Newcastle disease virus NP gene for nucleoprotein, genomic RNA,
DEFINITION strain 99299.
ACCESSION AJ306302
VERSION AJ306302.1 GI:28269508
KEYWORDS NP gene; nucleoprotein.
SOURCE Newcastle disease virus
ORGANISM Newcastle disease virus
VIRUSES; ssRNA negative-strand viruses; Mononegavirales;
Paramyxoviridae; Paramyxovirinae; Rubulavirus.
REFERENCE 1
Barbezange, C. and Jestin, V.
Molecular characterisation of three avian paramyxovirus type 1
isolated from pigeons in France
JOURNAL Virus Genes 26 (2), 175-183 (2003)
MEDLINE 22687690
PUBMED 12803469
REFERENCE 2 (bases 1 to 1580)
Barbezange, C.
Direct Submission
AUTHORS Submitted (07-FEB-2001) Barbezange C., VIPAC, AFSSA Site de
JOURNAL Ploufragan, BP53, 22440 Ploufragan cedex, FRANCE
FEATURES
source
location/Qualifiers
1..1580
/organism="Newcastle disease virus"
/virus
/mol_type="genomic RNA"
/strain="99299"
/specific_host="pigeon"
/db_xref="taxon:11176"
/country="France"
/note="pigeon paramyxovirus type 1"
14..1483
/gene="NP"
14..1483
/gene="NP"
/codon_start=1
```

```
/product="nucleoprotein"
/protein_id="CAC83752.1"
/db_xref="GI:28269509"
/translation="MSSVFDEYEQLLAAQTRPNGAHGGEGKSTLKVPPVPTLNSDD
PIDSWTSVFCRLRIAVSDEDANKPLRQGLAISLLGSHSOVMNRNHALACAKONEATLAVL
BEDFTNSVPOFNRSRGSVEERAFRPMVIAESLPACSNNTPFVTAGVENDVPEITD
TLERLSIAQAVWTVAKAMTAYETADESETRINKYMQQSRVKKYLLHPPVCSAIQ
LTTHSLAVRIFLSELKRGNTAGGSTYYNLVDVDSYIRNTGLTFAFTLYLKYGIN
TKTSALASLSDIQQKQMLRYMRKGENAPYMTLLGSDQMSFAFAEYLYSFA
MMAASVLDKGTGKQFARDFMSTFMRIGVEYAAQSSINEDMAAEKLKLPAAARRGL
AAAAQVSEETGNDIPTQOAGVLTLGSDKGPRAQGGPSRSQGPADAGDGETQFLDL
MRVANSNMRRAQSDSIHPEPLSTHPSQDNDTDMGY"
```

## ORIGIN

```
Query Match 86.1%; Score 1265.2; DB 14; Length 1580;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
QY 1 ATGTCTTCCTCGTATTCGATGAATACGAGCAGCTCCTCGCTGCTCAGACTCGCCCCCATGGA 60
Db 14 ATGTCTTCCTCGTATTCGATGAATACGAGCAGCTCCTCGCTGCTCAGACTCGCCCCCATGGA 73
QY 61 GCTCAGGAGGGGAGAGAGAGGAGGAGCATTAAAGAGTTGAGGTCCTCAGTATTCCTCTT 120
Db 74 GCTCAGGAGGGGAGAGAGAGGAGGAGCATTAAAGAGTTGAGGTCCTCAGTATTCCTCTT 133
QY 121 AACGTGACGATCCAGAGATAGATGGAATTTTGGGTATTTCTGTCTTCGAGTTGCTGTT 180
Db 134 AACGTGATGATCCAGAGATAGATGGAATTTTGGGTATTTCTGTCTTCGAGTTGCTGTT 193
QY 181 AGCGAGGAGCGCAACAAACCGCTCAGGCAAGGTCCTCATATCCCTCTCTGCTCCCAT 240
Db 194 AGTGAAGATGATCCAGAGATAGATGGAATTTTGGGTATTTCTGTCTTCGAGTTGCTGTT 253
QY 241 TCTCAAGTATGAGGAAACCATGTTGCCCTTCAGGAAACAGAAATGAGGCTACACTGACT 300
Db 254 TCTCAAGTATGAGGAAACCATGTTGCCCTTCAGGAAACAGAAATGAGGCTACACTGACT 313
QY 301 GTTCTTGAGATCGATGTTTACAGCAGCGTGCTCAGTTCAAACAAGAGTGGGGTG 360
Db 314 GTTCTTGAGATCGACAGTGTTCACCAAGTGTGCCCAAGTTCAACAAGAGCGGAGTG 373
QY 361 TCTGAGGAGGAGCAGCAGAGATTCATGGTGTATGACAGGGTCTCTCCCTCGGGGCTGAGT 420
Db 374 TCTGAAGAGGAGCAGCAGAGATTCATGGTGTATGACAGAGTCTCTCCCGGGGCTGAGT 433
QY 421 AACGGTACTCCGTTCTGTCAGCGCTGGGTTGAAGATGATGACCCAGAGAGATATCACTGAT 480
Db 434 AACAACTACTCGTTCTGTTACAGCTGGAGTTGAANTGATGTCAGAGAGATATCACTGAC 493
QY 481 ACTCTGGAAGAATCTCTGTTATCCAGGCTCAGGTATGGGTCAAGTAGCGAAGGCCATG 540
Db 494 ACTCTGGAAGAATCTCTGTTATCCAGGCTCAGGTATGGGTCAAGTAGCGAAGGCCATG 553
QY 541 ACTGCATATGAGCAGCAGATGAGTCCGGAACAGAGAGATCAATAGTATGATCAGCA 600
Db 554 ACTGCATATGAGCAGCAGATGAGTCCGGAACAGAGAGATCAATAGTATGATCAGCA 613
QY 601 GGCAAGTCCAGAGAAGATACATCCTCACCTGTATGACAGGAGTCAATTCACCTCACA 660
Db 614 GGTAGATCCAGAGAAGATACATCCTCACCTGTATGACAGGAGTCAATTCACCTCACA 673
QY 661 ATCAGACATTTCTGCGAGTCCGCAATTTCTTGTAGTACGAGCTTAAGAGAGGCGCAAT 720
Db 674 ATCAGACATTTCTGCGAGTCCGCAATTTCTTGTAGTACGAGCTTAAGAGAGGCGCGCAAT 733
QY 721 ACAGGAGGTTGGAGTCCAGATTCACAACTTGTAGTGGGATGTAGACTATCATCAGG 780
Db 734 ACAGGAGGTTGGAGTCCAGATTCACAACTTGTAGTGGGATGTAGACTATCATCAGG 793
QY 781 AACACCGGCTTACTGCTGATTTCTTCTTACACTCAATATGAAATTAATACCAACATCA 840
Db 794 AACACCGGCTTACTGCTGATTTCTTCTTACACTCAATATGAAATTAATACCAACATCA 853
```

QY	841	GCCTAGCACTCAGCAGCTCTACAGCGATATCCAAAGATGAAGCAGCTCATGCGTTTA	900
Db	854	GCCTTGCACTCAGCAGCTCTCAGGTATATCCAAAATGAACAGCTCATGCGTTTA	913
QY	901	TATCGATGAAGGAGAAATCGCGTACATGACATTCCTAGGTGACAGTATCAGATG	960
Db	914	TATCGATGAAGGAGAAATGCAATGATACATTCCTAGGTGACAGTATCAGATG	973
QY	961	AGCTTTGACCGGCTGAGTATGACAGCTTTATTTCTTTGCGATGGCATGATGATC	1020
Db	974	AGCTTTGACCGGCTGAGTATGACAGCTTTATTTCTTTGCGATGGCATGATGATC	1033
QY	1021	TTAGATAAGGAACTGGCAAAATACCAATTCGCCAGAGCTTCATGAGCACATCTTCG	1080
Db	1034	TTAGATAAGGAACTGGCAAAATACCAATTCGCCAGAGCTTCATGAGCACATCTTCG	1093
QY	1081	AGACTCGGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAGACATGCTGCT	1140
Db	1094	AGACTCGGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAATGAGGATATGCTGCT	1153
QY	1141	GAGCTAAACTAACCCCGGACAGAGAGGGGCTGGAGCTGCTGCCCAACAGAGTCT	1200
Db	1154	GAGCTAAACTAACCCCGGACAGAGAGGGGCTGGAGCTGCTGCCCAACAGAGTCT	1213
QY	1201	GAGGAACTGGCAGCTGATATTCCTACTCAACAGCGGGGTCTCACTGGGCTCAGC	1260
Db	1214	GAGGAACTGGCAGCTGATATTCCTACTCAACAGCGGGGTCTCACTGGGCTCAGC	1273
QY	1261	GATGAGGCCCCCGAGCTCTCAGGCTGATGCAACCAAGTCGCAAGGGCAACAGATGCC	1320
Db	1274	GACAAAGTCCCGAGCTCCGAGGCTGACCGGAGCAGTTCGCAAGGGCAACCGATGCT	1333
QY	1321	GGAGATGGGAGACCAATCTTGATTTGATGAGAGAGTGGCGAAGCATGCGAGAA	1380
Db	1334	GGGAGATGGGAGACCAATCTTGACCTGATGAGAGAGTGGCGAAGCATGCGAGAA	1393
QY	1381	GGCCAACTCCGACAGAGCACCCACCCCGGAAACCCCGGACTCCCGGGCCATCA	1440
Db	1394	GGCCAGATTCGACAGAGCACCATTCACCCGGAGCCCTCTCGACTCATGAGCCATC	1453
QY	1441	CAAGATAACGACACCGACTGGGGTATTGA	1470
Db	1454	CAGGACACGACACCGACTGGGGTACTGA	1483
RESULT 9	AF419407	1488 bp RNA linear	VRL 21-FEB-2002
LOCUS	NEWCASTLE DISEASE VIRUS PHEASANT/U.S.(CA)/F98-1208/98	nucleocapsid protein gene, complete cds.	
DEFINITION	AF419407		
ACCESSION	AF419407.1	GI:15811681	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			

1. .1488

/organism="Newcastle disease virus"

/mol\_type="genomic RNA"

/isolate="pheasant/U.S.(CA)/F98-1208/98"

/db\_xref="taxon:111176"

4. .1473

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAL09060.1"

/db\_xref="GI:15811682"

/translation="MSSVFDEYEQLLAAQTRPNHTGGGKSTLKVVPVFTLNSDD  
PEDRNFAVCLRFVAVSDANKPQGLALISLCSHVSQVMRNVHVALAGKQNDATLAVL  
EIDGFTNNVPOFSNRSRGSSEBAORFVIAAGSLPRACSNPTFVTVAGVDEPEIDITD  
TLERILSTQAVVTVAKAMTAYETADESETINKYMOQGRVQKYLHPVCRGAIO  
LTIHSIAVRIFLYSELKGRNTAGGSSYYNLVGDVSYIRNTGLTFAFFLTKYGIN  
TKTSALAUSSITGDIQKMKLMRLYRMKGENAPYMLLGDSDQMSFAFAEYALISFA  
MGMAVLDKGTGYQFARDFMSTSFWRIGVEYAOQGSINEDMAELKLTTPAARGL  
AAAQSVSEETGSMIDPTQAGVLTLGLDEBETPQSGSNKPGQPDAGDGTQFLDF  
MRAVANSRMREAPSPAQSTTHPEPPPTPGPPQDDDDTWGY"

Query Match

Best Local Similarity

Matches 1337; Conservative

0; Mismatches

133; Indels

0; Gaps

0;

QY

1

ATGTCTTCGTTATTCGATGATACGAGCAGCTCTCGCTGCTCAGACTCGCCCCAATGA

60

Db

4

ATGTCTTCGTTATTCGATGATACGAGCAGCTCTCGCTGCTCAGACTCGCCCCAATGA

63

QY

61

GCTCAGGAGGGGAGAGAGGGAGCATTAAAGAGTTGAGTCCAGTATTCATCTT

120

Db

64

ACTCAGGAGGGAGAGAGGGAGCATTAAAGAGTTGAGTCCAGTATTCATCTT

123

QY

121

AACAGTACGATCCAGAGATAGATGGAATTTTCGGTATTTCTGCTTCGGATTCGTT

180

Db

124

AACAGTACGATCCAGAGATAGATGGAATTTTCGGTATTTCTGCTTCGGTTCGTT

183

QY

181

AGCAGGAGCGCCAAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGCTCCCAT

240

Db

184

AGTGAAGTCCCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGCTCCCAT

243

QY

241

TCTCAAGTATGAGGAAACCATTTGCCCTTGAGGAAACAGAAATGAGGCTACACTGAT

300

Db

244

TCTCAGGTGATGAGAAACCATTTGCCCTTGAGGAAACAGAAATGAGGCTACACTGAT

303

QY

301

GTTCTTGAGATCGATGGTTTACAGCAGCGTCCCTCAGTTCACCAAGAGGAGTGGGTG

360

Db

304

GTTCTTGAGATCGATGGTTTACAGCAGCGTCCCTCAGTTCACCAAGAGGAGTGGGTG

363

QY

361

TCTGAGGAGAGCAGACAGAGATTCATGGTATGAGCAGGGTCTCTCCCTCGGGCGTCACT

420

Db

364

TCTGAGGAGAGCAGACAGAGATTCATGGTATGAGCAGGGTCTCTCCCTCGGGCGTCACT

423

QY

421

AACGGTACTCCGTTTCGTCACGGCTGGGTTGAAGATGATGACCCAGAGATATCACTGAT

480

Db

424

AACGGTACTCCGTTTCGTCACAGCTGGGTTGAAGATGATGACCCAGAGATATCACTGAC

483

QY

481

ACTCTGGAAGATCTCTGCTATCCAGGCTCAGGTATGGTTCAGTACCGAGGCGCATG

540

Db

484

ACCTTGGAAGATCTCTGCTATCCATCCAGCTCAGGTATGGTTCAGTACCGAGGCGCATG

543

QY

541

ACTGCTATGACAGCAGCAGATGAGTGGGTTGAAGATGATGACCCAGAGATATCACTGAC

600

Db

544

ACTGCTATGACAGCAGCAGATGAGTGGGTTGAAGATGATGACCCAGAGATATCACTGAC

603

QY

601

GGCAGAGTCCAGAGAGTATCTCTCCACCTGTATGAGGAGTGCATTTCACTCACA

660

Db

604

GGTAGAGTTTCAAGAGAGTATCTCTTCACTGTATGAGGAGTGCATTTCACTCACA

663

QY

661

ATCAGACATTTCTGCGAGTCCGCAATTTCTTAGTTAGGAGCTTAAGAGAGCGCGCAAT

720

Db

664

ATCAGACATTTCTGCGAGTCCGCAATTTCTTAGTTAGGAGCTTAAGAGAGCGCGCAAC

723

Qy	721	ACGCGAGGTGGGAGCTCCACGTATTACAATCTAGTAGGGGATGTAGACTCATCATCAGG	780
Db	724	ACAGCAGTGGGAGCTCCACATATTACAATCTAGTCGGGATGTAGACTCATCATCAGA	783
Qy	781	AACACGGAGCTTACTGCATTTCTTACACTCAAAATATGAAATTAATACCAAGACATCA	840
Db	784	AACACGGGCTTACTGCATTTCTTACACTCAAAATATGAAATTAATACCAAGACGTCA	843
Qy	841	GCCCTAGCACTCAGCAGCCTCACAGGCGATATCCAAAAGATGAAGCAGCTCATCGCGTTTA	900
Db	844	GCCCTTGCACTCAGCAGCCTCACAGGTGATATCCAAAAATGAACAACATCATGCGTTTA	903
Qy	901	TATCGGATGAAGGAGAAAATGCGCCGTATCATGACATGTCTAGGTGACAGTATCAGATG	960
Db	904	TATCGGATGAAGGTGAAAATGCAACCGTATATGACATTTGCTAGGTGACAGTGACCAAGT	963
Qy	961	AGCTTTGCACCGCTGAGTAGTGCACAGCTTTATTTCTTTCCCATGGGCATGGCATCAGTC	1020
Db	964	AGCTTTCGCAACAGCTGAGTAGTGCACAACTTTATTTCTTTGCCATGGGCATGGCGTCAGTC	1023
Qy	1021	TTAGATAAAGGAACTGGCCAAATACCAATTCGCCAGAGACTTTCATGAGCACATCATTTCTGG	1080
Db	1024	TTAGATAAGGGAATGGCAAGTACCAATTCGCCAGGCACTTCATGAGCACATCATTTCTGG	1083
Qy	1081	AGACTCGGGTGGGATATGCTCAGGCTCAGGGGAGTAGCATCAACGAAGACATGGCTGCT	1140
Db	1084	AGACTTGGAGTAGAGTATGCTCAGGCTCAAGGAAGTAGCATTAACGAGGACATGGCTGCT	1143
Qy	1141	GAGCTAAACTTAACCCGGCAGCAAGAGGGGCTTCGACGCTGTCGCCAACGATGCTCT	1200
Db	1144	GAGCTGAAACTTAACCCAGCAGCAAGGAGGGCTTCGCAGCTGCTGCCAACGAGTATCT	1203
Qy	1201	GAGGAAACTTGGCAGCTGGATATTTCTTACTCAACGAAGCCGGGTTCCTCATCTGGGCTCAGC	1260
Db	1204	GAGGAAATCGCAGCATGGACATTTCTTACTCAACAGGGGGAGTCTCTCACGGGCTCAGT	1263
Qy	1261	GATGGAGGCCCCGAGCTCTCTCAGGTTGGATTCGAACAAAGTCGAAGGGCAACAGATGCC	1320
Db	1264	GACGAAGCCCCGAAACCCACAGAGCGGATCAAAACAAGCCGAAGGGCAACAGATGCC	1323
Qy	1321	GGAGATGGGAGACCCAAATTTCTTGGATTTGATGAGAGCAGTGGGGAACAGCATGCGAGAA	1380
Db	1324	GGGAGATGGGAGACTCAATTTCTTGGATTTTATGAGAGCAGTGGGCGAATGATCATGCGAGAG	1383
Qy	1381	GGCGCAAACTCCGACAGAGCACCAACCCCGGAACCCCGCCGACTCCCGGGGCCATCA	1440
Db	1384	GGCGCAAGTCCCGACAGAGGACACCAACCATCCGAGAGCTCCCCCACTCTCTGGGCCACCA	1443
Qy	1441	CAAGATAACGACACCGACTCGGGGTATTGA	1470
Db	1444	CAAGACGATGACACTGACTGGGGGTACTAA	1473

RESULT 10	AF419396	1585 bp	RNA	linear	VRL 21-FEB-2002
LOCUS	AF419396				
DEFINITION	Newcastle disease virus parrot/U.S.(IL)/27994/91 nucleocapsid protein gene, complete cds.				
ACCESSION	AF419396				
VERSION	AF419396.1	GI:15811659			
KEYWORDS	.				
SOURCE	Newcastle disease virus				
ORGANISM	Newcastle disease virus				
REFERENCE	Viruses; ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Paramyxovirinae; Rubulavirus.				
AUTHORS	1 (bases 1 to 1585)				
TITLE	Seal,B.S., Crawford,J.M., Sellers,H.S., Locke,D.P. and King,D.J. Nucleotide sequence analysis of the Newcastle disease virus nucleocapsid protein gene and phylogenetic relationships among the Paramyxoviridae				
JOURNAL	Virus Res. 83 (1-2), 119-129 (2002)				
MEDLINE	21854562				
PUBMED	11864745				

REFERENCE	2 (bases 1 to 1585)
AUTHORS	Seal,B.S., Crawford,J.M., Sellers,H.S., Locke,D.P. and King,D.J.
TITLE	Direct Submission
JOURNAL	Submitted (14-SEP-2001) Agricultural Research Service, USDA, Southeast Poultry Research Laboratory, 934 College Station Road, Athens, GA 30605, USA
FEATURES	Location/Qualifiers
source	1. .1585
	/organism="Newcastle disease virus"
	/mol_type="genomic RNA"
	/isolate="parrot/U.S.(IL)/27994/91"
	/db_xref="taxon:11176"
CDS	67. .1536
	/codon_start=1
	/product="nucleocapsid protein"
	/protein_id="A1509049.1"
	/db_xref="GI:15811660"
	/translation="MSVSFDEYQLLAAQTRPNGAHGGEGKSTLKVYVVPVPIINSDD PEIDWNPVSFLKRIAVSDEDKPLRGALISLCSHVSQNMNHVALAGKQNEATLAIL EDIGFANSPTQFNRRSGVSERAPFMWJTAGSLPRACNSNGTFFVTAGVEDADEPITD TLRIILSIAQVWVTVAKAMTAYETADESETRINKYMQQVRQVKYLIHPVCRSAIQ LTIRLSLAVRI FLVSELRGRNTAGSSTVYLVGVDVSYIRNTGLTAPFLTKLYGIN TKTISALSLSDGI QMKQLMELRYMKGENAPYMTLLGDSQMSFAPAEYAOIYSPA MMAWSVLDTGKTKYQFAPDMTSTFWRLGVEYAOAGSSINEDMAELKLTTPAARGL AGGASVEIGSVSDVPTQGMVLTGLSDVGPOAPQGSNKSXGPDAGDGETQFLDM MRVAVNSMREAQNSAFSTPQGPPEPTPGPSQNNDTWGY"
ORIGIN	
Query Match	85.5%; Score 1257.2; DB 14; Length 1585;
Best Local Similarity	91.0%; Pred. No. 0;
Matches 133;	Conservative 0; Mismatches 133; Indels 0; Gaps 0;
Qy	1 ATGTCTTCCGTATTTCGATGAATACGAGCAGCTCTCGCTGCTCAGACTCGCCCAATGGA 60
Db	
Qy	67 ATGTCTGTCGTATTTGACGATACGACCACTCTCTGCTGCTCAGCCCGCCCAATGGA 126
Db	
Qy	61 GCTCACGGAGGGGAGAGAGAGGAGGACATTTAAGAGTTGAGGTCCAGATTTCACCTTT 120
Db	
Qy	127 GCTCACGGGGGAGAGAGAGGGGAGCACTTTGAAGTTGAGGTCCAGATTATTACTCTT 186
Db	
Qy	121 AACAGTCACATCCAGAAGATAGATGGAATTTTCGGGTATTTCTGTCTTCGGATGCTGTT 180
Db	
Qy	187 AATAGTCACGATCCAGAAGACAGATGGAATTTTTCGGTATTTCTGTCTTCGGATTGCCGTT 246
Db	
Qy	181 AGCGAGACGCCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGTGCTCCCAT 240
Db	
Qy	247 AGCGAGATGCCAACAGCCACTCAGSCAAGGTGCTCTTATATCCCTCTTGTGCTCCAC 306
Db	
Qy	241 TCTCAAGTGATGAGGAACCATGTTGCCCTTTCGAGGAAAAACAGATAGGGCTACACTGACT 300
Db	
Qy	307 TCTCAAGTGATGAGGAACCATGTTGCCCTTTCGAGGAAAAACAGATAGGGCCACACTAGCT 366
Db	
Qy	301 GTTCTTGAGATCGATGTTTTTACCAGCAGCGTGCCTCAGTTTCAACACAGAGTGGGGTG 360
Db	
Qy	367 ATTCTTGAGATTGATGGGTTTGGCAACAGACGCCCTCAGTTTCAACACCGGAGCGGAGTG 426
Db	
Qy	361 TCTGAGCAGAGAGCACAGAGATTATCGTGATAGCAGGGTCTCTCCCTCGGGCGTGCACT 420
Db	
Qy	427 TCGAGAGAGAGGACACAGAGATTATGGTGATAGCCGGGTCTCTCCCTCGGGCATGCGC 486
Db	
Qy	421 AACGGTACTCCGTTTCGTTCACGGCTGGGGTTGAAGATGATCACCCAGAAAGATATCACTGAT 480
Db	
Qy	487 AACGGTACCCGTTTCGTTCACAGCCGGGTTTGAGATGACGACCCAGAAACATTTACTGAC 546
Db	
Qy	481 ACTCTGAAAGAATTCCTGTCTATCCAGGCTCAGGTATGGGTACAGTAGGGAAGGCCATG 540
Db	
Qy	547 ACTCTGAAAGGATCCTTATCTATCCAGGCTCAGGTATGGGTACAGTAGGGAAGGCTATG 606
Db	
Qy	541 ACTGCATATCAGACAGAGATGATCGGAACACAGAGAGATCAATAGTACATGCAGCAA 600
Db	
Qy	607 ACTGCATATGAACACAGAGATGATCGGAACACAGAGAGATCAATAGTATATGACAGCAA 666
Db	
Qy	601 GGCAGAGTCCAGAAAGATGATCATCTCTCCACCTGTATGCAGGAGTGCATTTCAACTCACA 660
Db	

Db	667	GCACAGTCCAAAGAAATACATCTCCACCCGATGACAGAGTGCAATTCATCTCA	726
QY	661	ATCAGACATCTCTGGAGTCGCGATTTCTTAGTTAGCGAGCTTAAGAGAGCGCGCAAT	720
Db	727	ATCAGGCAATCTTTGGCAGTTCCGATTTCTTAGTTAGCGAGCTTAAGAGAGCGCGCAAT	786
QY	721	ACGGCAGTGGGAGCTCCACGTAATACAACTTAGTAGGGGATGAGCTCATACATCAGG	780
Db	787	ACGGCAGTGGGAGCTCCACATATTACAACTTTGGTAGGGGATGAGCTCATATATCAGA	846
QY	781	AACACCGGACTTACTGCAATTTCTTCTTACATCAAAATATGGAATTAATACCAAGACATCA	840
Db	847	AACACTGGGCTTACTGCATTTCTCTGACATCAATATGGATTAATACCAAGACATCA	906
QY	841	GGCCTAGCACTCAGCAGCGCTCA CAGCGGATATCCAAAAGATCAAGCAGCTCATGCTTTA	900
Db	907	GCTCTTGCACTCAGCAGCGCTCA CAGGTGATATCCAAAAATTAAGCAGCTCATGCTTTA	966
QY	901	TATCGGATGAAGGGAGAAAATCGCGGTACATGACATTCGTAAGTGACAGCTGATCAGATG	960
Db	967	TATCGGATGAAGGGAGAAAATCGCGGTACATGACATTCGTAAGTGACAGCTGATCAGATG	1026
QY	961	AGCTTTGCA CCGGCTGAGTATGCA CAGCTTTATTTCTTGCCATGGCATGCATCAGTC	1020
Db	1027	AGCTTTGCA CCGGCTGAGTATGCA CAACTTTATTTCTTGCCATGGCATGCATCAGTC	1086
QY	1021	TTAGATAAAGGAATGGGCAAAATACCAATTCGCCAGAGCTTCATGAGCACATCATCTCGG	1080
Db	1087	TTAGATAAAGGAATGGGCAAAATACCAATTCGCCAGAGCTTCATGAGCACATCATCTCGG	1146
QY	1081	AGACTCGGGTGGAGTATGCTCAGGCTCAGGGAGTATGATCAACAGAGACATGCTGCT	1140
Db	1147	AGGCTTGGGTGGAAATATGCTCAGGCTCAGGGAATGATCAATTAATGAGGATATGCTGCT	1206
QY	1141	GAGCTAAACTAACCCCGCAGCAAGAGGGGCTTGGCAGCTGCTGCCCAACGAGTGTCT	1200
Db	1207	GAGCTAAACTAACCCCGCAGCAAGAGGGGCTTGGCAGCTGCTGCCCAACGAGTGTCT	1266
QY	1201	GAGGAATCGCAGCGTGGATATTCCTACTCAACAGCGGGGTCTCATCTGGGCTCAGC	1260
Db	1267	GAGGAATTCGCGAGCGTGGATATTCCTACTCAACAGCGGGGTCTCATCTGGGCTCAGC	1326
QY	1261	GATGAGGCGCCCGAGCGCTCTCAGGTTGATCGAACAAGTCGCAAGGGGCAACAGATGCC	1320
Db	1327	GACGTAGGCGCCCGAGCGCTCTCAGGTTGATCGAACAAGTCGCAAGGGGCAACAGATGCC	1386
QY	1321	GGAGATGGGGAGACCCCAATTTCTTGGATTTGATGAGAGCAGTGGCGAACAGCATGCGAGAA	1380
Db	1387	GGGGACCGAGAGACCCAGTTCTTTGGATATGATGAGGCGGTGGCAAAATAGCATGCGAGAA	1446
QY	1381	GGCCCAAACTCGCACAGACACCCACCCACCGGAACCCCGGCTCTCCCGGGGCATCA	1440
Db	1447	GCACAAAATTCGACCGAGACCCCTCAACCGGGGCTCTCCCGGCTCTGGGGCTCT	1506
QY	1441	CAAGATAACGACACCGACTGGGGGTATTGA	1470
Db	1507	CAAAACATGACACCGACTGGGGGTACTGA	1536
RESULT 11			
AF419401			
LOCUS			
DEFINITION	AF419401	Newcastle disease virus chicken/Italy/1/00 nucleocapsid protein	1694 bp RNA linear VRL 21-FEB-2002
ACCESSION	AF419401	gene, complete cds.	
VERSION	AF419401.1	GI:15811669	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			

AUTHORS  
TITLEJOURNAL  
MEDLINE  
PUBMED

## REFERENCE

AUTHORS  
TITLE  
JOURNALFEATURES  
source

## CDS

## ORIGIN

Query Match	85.5%;	Score 1257.2;	DB 14;	Length 1694;
Best Local Similarity	91.0%;	Pred. No. 0;		
Matches 1337;	Conservative	0;	Mismatches 133;	Indels 0; Gaps 0;
QY	1	ATGCTTCCTCGTATTCGATGAATACGAGCAGCTCCCTCGCTGCTCAGACTCGCCCCCAATGA	60	
Db	15	ATGCTGCTGCTGCTTTTGTATGAATACGACAGCTCCCTCGCTGCTCAGACCGCCCTAACGGA	74	
QY	61	GCTCAGCAGGAGGGGAGAGAGGAGGAGCATTAAAGAGTTGAGTCCAGTATTCCTCTT	120	
Db	75	ACTCATGAGGAGGAGAGAGGAGGAGCATTAAAGAGTTGAGTCCAGTATTCCTCTT	134	
QY	121	AACAGTACGATCCAGAGAGATAGATGAATTTTTCGGGTATTCCTGCTTCGGATTCGCTGT	180	
Db	135	AACAGTACGATCCAGAGAGATAGATGAATTTTTCGGGTATTCCTGCTTCGGATTCGCTGT	194	
QY	181	AGCAGGAGCGCAACAAACCGCTCAGGCAAGTGTCTCTATATCCCTCCTGCTGCTCCCAT	240	
Db	195	AGCAGGAGCGCAACAAACCGCTCAGGCAAGTGTCTCTATATTCCTCTCTGCTCCCAT	254	
QY	241	TCTCAAGTGATGAGCAACCATGTTGCGCTTCGAGGAAACAGAGATGAGGCTACACTGACT	300	
Db	255	TCTCAGTGATGAGCAACCATGTTGCGCTTCGAGGAAACAGAGATGAGGCTACACTGACT	314	
QY	301	GTTCTTGAGATCGATGGTTTTTACAGCAGCGCTGCGCTCAGTTTCAACACAGGAGTGGGGTG	360	
Db	315	GTTCTTGAGATCGATGGTTTTTACAGCAGCGCTGCGCTCAGTTTCAACACAGGAGTGGGGTG	374	
QY	361	TCTGAGGAGAGCAGCAGAGATTCATGCTGATAGCAGGCTCTCTCCCTCGGCGCTGCTG	420	
Db	375	TCTGAGGAGAGCAGCAGAGATTCATGCTGATAGCAGGCTCTCTCCCTCGGCGCTGCTG	434	
QY	421	AACGCTACTCCTGCTCAGCGCTGGGGTTGAGAGATGATGACACAGAGATATCACTGAT	480	
Db	435	AACGCTACTCCTGCTCAGCGCTGGGGTTGAGAGATGATGACACAGAGATATCACTGAT	494	
QY	481	ACTCTGGAAGAAATCCTCTCTATCCAGGCTCAGGCTATGGGTACAGTACAGGAGGCGCATG	540	



Db 495 ACTTTGGAAGAATTCTATCTGTCCAGCTCAGGTATGGATCACAGTAGCAAAAGGCCATG 554  
 QY 541 ACTGCATATGACAGCAGATGAGTCGGAAACAGAAAGAAATCAATAGTATACGACAA 600  
 Db 555 ACTGCATATGACAGCAGATGAGTCGAAACAGAAAGAAATCAATAGTATATCGACAG 614  
 QY 601 GGCAGAGTCCAGAAAGATACATCTCCACCCCTGTATGCAGAGTGCATTAATCAACTACA 660  
 Db 615 GGGAGAAATTCAGAAAGAAATACATCTCCACCCCTGTATGCAGAGTGCATTAATCAACTACA 674  
 QY 661 ATCAGACATTTCTTGGCAGTCCGATTTCTTATAGTTAGCGAGCTTAAGAGAGGCGCAAT 720  
 Db 675 ATCAGGCAATTTCTTGGCAGTCCGATTTCTTATAGTTAGCGAGCTTAAGAGAGGCGCAAT 734  
 QY 721 ACAGCAGTGGAGTCCAGGATTTACAACTTATAGTGGGATGTAGACTCATATCATCAGG 780  
 Db 735 ACAGAGTGGAGTTCACATATTTACAACTTATAGTGGGATGTAGACTCATATCATCAGG 794  
 QY 781 AACACCGGACTTACTGCATTTCTTCTTACACTCAAAATATGAAATTAATACCAAGACATCA 840  
 Db 795 AACACCGGCTTACTGCATTTCTTCTTACACTCAAGTATGAAATCAATACCAAGACATCA 854  
 QY 841 GCCCTAGACTCAGCAGCTCAGAGCGATATCCAAAGATGAAAGAGTCAATCGCTTTA 900  
 Db 855 GCCCTTGCCTCAGCAGCTCAGAGCGATATCCAAAGATGAAAGAGTCAATCGCTTTA 914  
 QY 901 TATCGGATGAAGGAGAGAAATGCCCGTACATGATCTTCTAGCTGACAGTGCAGTGCAGT 960  
 Db 915 TATCGGATGAAGGAGAGAAATGCCCGTACATGATCTTCTAGCTGACAGTGCAGTGCAGT 974  
 QY 961 AGCTTTGACCGGCTGAGTATGCACAGCTTTATTTTGGCATGGGATGGCATCAGTC 1020  
 Db 975 AGCTTTGACCGCTGAGTATGCACAACTTTATTTCTTGGCATGGGATGGCATCAGTC 1034  
 QY 1021 TTAGATAAGAACTGGCAATATCAATTCGCCAGAGATTCATGAGCACATATTTCTGG 1080  
 Db 1035 TTGATAAGGAACTGGCAATATCAATTCGCCAGGATTTTATGAGCACATATTTCTGG 1094  
 QY 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGATGATCATCAAGAGATGGCTGCT 1140  
 Db 1095 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGATGATCATTAATGAGGACATGGCTGCT 1154  
 QY 1141 GAGCTAAACTAAACCCCGGAGCAAGAGGGGCTTGGCAGCTGTGCTGCCAACAGAGTGTCT 1200  
 Db 1155 GAGCTTAAGCTAACTCCCGGAGCAAGAGGGCTTGGCTGTGCTGCCAACAGAGTGTCT 1214  
 QY 1201 GAGGAACTGGCAGCTGGATATTCCTACTCAACAGCGGGGTCTTCACTTGGGCTCAGC 1260  
 Db 1215 GAGGAACTGGCAGCTGGATATTCCTACTCAACAGCGGAGGTCTTCACTTGGGCTTGGC 1274  
 QY 1261 GATGAGGCCCCGAGGCTCTCAGGGTGGATCGAAACAGTGCAGAGGCAACAGATGCC 1320  
 Db 1275 GACGAGGCCCCGAGGCTTCAAGGGGATCAACAGGCTTCAAGGGCAACAGATGCC 1334  
 QY 1321 GGAGATGGGAGACCCAAATTTCTTGGATTTGATGAGAGCAGTGGCGAAACAGCATGCGAGAA 1380  
 Db 1335 GGGAAATGGGAGACCCAAATTTCTTGGATTTGATGAGAGCAGTGGCGAAACAGCATGCGAGAA 1394  
 QY 1381 GCGCCAAATTCGCAAGAGACACCCACCGGAAACCCCGGAGTCTCCCGGAGTCCCGGAGTCA 1440  
 Db 1395 GCGCCAAATTCGCAAGAGACACCTGCCCATCCGAGCTCCCGGAGTCTCCCGGAGTCCCGGAGT 1454  
 QY 1441 CAAGATAACGACACCGGAGTGGGATTTGA 1470  
 Db 1455 CAAGATAACGACACTGACTGGGGTACTGA 1484

RESULT 12

AP419410

LOCUS

DEFINITION

Newcastle disease virus turkey/U.S. (ND)/43084/92 nucleocapsid

protein gene, complete cds.

AP419410

ACCESSION

## VERSION

AP419410.1 GI:15811687

## KEYWORDS

## SOURCE

## ORGANISM

Newcastle disease virus  
 Newcastle disease virus  
 Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 Paramyxoviridae; Paramyxovirinae; Rubulavirus.

## REFERENCE

## AUTHORS

## TITLE

Seal, B.S., Crawford, J.M., Sellers, H.S., Locke, D.P. and King, D.J.  
 Nucleotide sequence analysis of the Newcastle disease virus  
 nucleocapsid protein gene and phylogenetic relationships among the  
 Paramyxoviridae

Virus Res. 83 (1-2), 119-129 (2002)

## JOURNAL

## MEDLINE

## PUBMED

2 (bases 1 to 1561)  
 11864745

## REFERENCE

Seal, B.S., Crawford, J.M., Sellers, H.S., Locke, D.P. and King, D.J.  
 Direct Submission

## AUTHORS

## TITLE

## JOURNAL

Submitted (14-SEP-2001) Agricultural Research Service, USDA,  
 Southeast Poultry Research Laboratory, 934 College Station Road,  
 Athens, GA 30605, USA

## FEATURES

location/Qualifiers

1..1561

/organism="Newcastle disease virus"

/mol\_type="genomic RNA"

/isolate="turkey/U.S. (ND)/43084/92"

/db\_xref="taxon:11176"

67..1536

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAU09063.1"

/db\_xref="GI:15811688"

/translation="MSSVFDEYEQLLAAQTRPNHAGGGKSTLKVFPVFTLNSDD  
 PEDRNFSVFLRIAVSEDEKPLRQALISLCSHSQVNRNHNVALAGKQNEATLAIL  
 EIDFTFTQPFNNRSGVSEERAFMVIAGSLPRACNSGTPFATGVEDDAPEDITD  
 TLRIILSIQAVVTVAKAMTAYETADESETRINKYQGGVQKKIIPHPVCRSAIQ  
 LTRSLAVRIFLFSELKGRNTTGGSTYYNLVGDVSVIRNTGLTAFELTKYGIN  
 KTTSALSLSLTGDIOKMKOLMBLYRMKGENAPYTWLLGSDSDQMSFAPAYQLYSFA  
 MGNASVLEKGTGKYQFARDEFMSTSFNRLGVEYQAQSSINEDMAELKUTPAARRGL  
 AAAQVRSEIGSMIDPTQAGVLTLGRDGGPQAPQGGSNKSGQPGAGTQFLDL  
 MRVANSREAPNSARSTTHPEPPTSGSPSPDNDTDWGY"

/db\_xref="taxon:11176"

67..1536

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAU09063.1"

/db\_xref="GI:15811688"

/translation="MSSVFDEYEQLLAAQTRPNHAGGGKSTLKVFPVFTLNSDD  
 PEDRNFSVFLRIAVSEDEKPLRQALISLCSHSQVNRNHNVALAGKQNEATLAIL  
 EIDFTFTQPFNNRSGVSEERAFMVIAGSLPRACNSGTPFATGVEDDAPEDITD  
 TLRIILSIQAVVTVAKAMTAYETADESETRINKYQGGVQKKIIPHPVCRSAIQ  
 LTRSLAVRIFLFSELKGRNTTGGSTYYNLVGDVSVIRNTGLTAFELTKYGIN  
 KTTSALSLSLTGDIOKMKOLMBLYRMKGENAPYTWLLGSDSDQMSFAPAYQLYSFA  
 MGNASVLEKGTGKYQFARDEFMSTSFNRLGVEYQAQSSINEDMAELKUTPAARRGL  
 AAAQVRSEIGSMIDPTQAGVLTLGRDGGPQAPQGGSNKSGQPGAGTQFLDL  
 MRVANSREAPNSARSTTHPEPPTSGSPSPDNDTDWGY"

/db\_xref="taxon:11176"

67..1536

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAU09063.1"

/db\_xref="GI:15811688"

/translation="MSSVFDEYEQLLAAQTRPNHAGGGKSTLKVFPVFTLNSDD  
 PEDRNFSVFLRIAVSEDEKPLRQALISLCSHSQVNRNHNVALAGKQNEATLAIL  
 EIDFTFTQPFNNRSGVSEERAFMVIAGSLPRACNSGTPFATGVEDDAPEDITD  
 TLRIILSIQAVVTVAKAMTAYETADESETRINKYQGGVQKKIIPHPVCRSAIQ  
 LTRSLAVRIFLFSELKGRNTTGGSTYYNLVGDVSVIRNTGLTAFELTKYGIN  
 KTTSALSLSLTGDIOKMKOLMBLYRMKGENAPYTWLLGSDSDQMSFAPAYQLYSFA  
 MGNASVLEKGTGKYQFARDEFMSTSFNRLGVEYQAQSSINEDMAELKUTPAARRGL  
 AAAQVRSEIGSMIDPTQAGVLTLGRDGGPQAPQGGSNKSGQPGAGTQFLDL  
 MRVANSREAPNSARSTTHPEPPTSGSPSPDNDTDWGY"

/db\_xref="taxon:11176"

67..1536

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAU09063.1"

/db\_xref="GI:15811688"

/translation="MSSVFDEYEQLLAAQTRPNHAGGGKSTLKVFPVFTLNSDD  
 PEDRNFSVFLRIAVSEDEKPLRQALISLCSHSQVNRNHNVALAGKQNEATLAIL  
 EIDFTFTQPFNNRSGVSEERAFMVIAGSLPRACNSGTPFATGVEDDAPEDITD  
 TLRIILSIQAVVTVAKAMTAYETADESETRINKYQGGVQKKIIPHPVCRSAIQ  
 LTRSLAVRIFLFSELKGRNTTGGSTYYNLVGDVSVIRNTGLTAFELTKYGIN  
 KTTSALSLSLTGDIOKMKOLMBLYRMKGENAPYTWLLGSDSDQMSFAPAYQLYSFA  
 MGNASVLEKGTGKYQFARDEFMSTSFNRLGVEYQAQSSINEDMAELKUTPAARRGL  
 AAAQVRSEIGSMIDPTQAGVLTLGRDGGPQAPQGGSNKSGQPGAGTQFLDL  
 MRVANSREAPNSARSTTHPEPPTSGSPSPDNDTDWGY"

/db\_xref="taxon:11176"

67..1536

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAU09063.1"

/db\_xref="GI:15811688"

/translation="MSSVFDEYEQLLAAQTRPNHAGGGKSTLKVFPVFTLNSDD  
 PEDRNFSVFLRIAVSEDEKPLRQALISLCSHSQVNRNHNVALAGKQNEATLAIL  
 EIDFTFTQPFNNRSGVSEERAFMVIAGSLPRACNSGTPFATGVEDDAPEDITD  
 TLRIILSIQAVVTVAKAMTAYETADESETRINKYQGGVQKKIIPHPVCRSAIQ  
 LTRSLAVRIFLFSELKGRNTTGGSTYYNLVGDVSVIRNTGLTAFELTKYGIN  
 KTTSALSLSLTGDIOKMKOLMBLYRMKGENAPYTWLLGSDSDQMSFAPAYQLYSFA  
 MGNASVLEKGTGKYQFARDEFMSTSFNRLGVEYQAQSSINEDMAELKUTPAARRGL  
 AAAQVRSEIGSMIDPTQAGVLTLGRDGGPQAPQGGSNKSGQPGAGTQFLDL  
 MRVANSREAPNSARSTTHPEPPTSGSPSPDNDTDWGY"

/db\_xref="taxon:11176"

67..1536

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAU09063.1"

/db\_xref="GI:15811688"

/translation="MSSVFDEYEQLLAAQTRPNHAGGGKSTLKVFPVFTLNSDD  
 PEDRNFSVFLRIAVSEDEKPLRQALISLCSHSQVNRNHNVALAGKQNEATLAIL  
 EIDFTFTQPFNNRSGVSEERAFMVIAGSLPRACNSGTPFATGVEDDAPEDITD  
 TLRIILSIQAVVTVAKAMTAYETADESETRINKYQGGVQKKIIPHPVCRSAIQ  
 LTRSLAVRIFLFSELKGRNTTGGSTYYNLVGDVSVIRNTGLTAFELTKYGIN  
 KTTSALSLSLTGDIOKMKOLMBLYRMKGENAPYTWLLGSDSDQMSFAPAYQLYSFA  
 MGNASVLEKGTGKYQFARDEFMSTSFNRLGVEYQAQSSINEDMAELKUTPAARRGL  
 AAAQVRSEIGSMIDPTQAGVLTLGRDGGPQAPQGGSNKSGQPGAGTQFLDL  
 MRVANSREAPNSARSTTHPEPPTSGSPSPDNDTDWGY"

/db\_xref="taxon:11176"

67..1536

/codon\_start=1

/product="nucleocapsid protein"

/protein\_id="AAU09063.1"

/db\_xref="GI:15811688"

## ORIGIN

## Query Match

85.2%; Score 1252.4; DB 14; Length 1561;

Best Local Similarity 90.7%; Pred. No. 0;

Matches 1334; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

Qy 1 ATGTCTTCCTGATTCGATGAATACGACAGCTCCTCGCTGCTCAGACTCGCCCCAATGGA 60

Db 67 ATGTCTGCTGATTTGACGAATACGACAGCACTCTCTGCTGCTCAGACTCGCCCCAATGGA 126

Qy 61 GCTCAGGAGGGGAGAGAGAGGAGGACATTTAAGATTGAGGTCCAGGTCCACATCTT 120

Db 127 GCTCATGAGGAGGAGAGAGGAGGACATTTAAGATTGAGGTCCAGGTCCACATCTT 186

Qy 121 AACAGTCAAGCATCAGAAAGATAGATGAAATTTTCGGGTATTTCTGCTTCCGATTCGCTT 180

Db 187 AATAGTCAAGCATCAGAAAGATAGATGAAATTTTCGGGTATTTCTGCTTCCGATTCGCTT 246

Qy 181 AGCGAGGAGCGCCAAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTCTGCTGCCAT 240

Db 247 AGCGAGGATGCCAAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTCTGCTGCCAT 306

Qy 241 TCTCAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300

Db 307 TCTCAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366

Qy 301 GTTCTTGGATTCGATGCTGTTTACCAGCAGCTGCTCAGTTCAGTTCAGCAGGAGTGGGTG 360

Db 367 ATTCTTGGATTCGATGCTGTTTACCAGCAGCTGCTCAGTTCAGTTCAGCAGGAGTGGGTG 426

Qy 361 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420

Db 427 TCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 486

```
421 AACGCTACTCCGTTCCGTCAACGGCTGGGGTTCAGAGTATGATGCACAGAGCATATCACTCAT 480
Db      |||
487 AACGCTACTCCGTTCCGTCAACGGCTGGGGTTCAGAGTATGATGCACAGAGCATATCACTCAT 546
Qy      |||
481 ACTCTGAAAGAAATCCTGTCTATCCAGGCTCAGGCTATGGGTACACAGTACGAGGCCCATG 540
Db      |||
547 ACTCTGAAAGAAATCCTGTCTATCCAGGCTCAGGCTATGGGTACACAGTACGAGGCCCATG 606
Qy      |||
541 ACTGCTATGACAGCAGAGTATGAGTACGAGAAACAGAGAAATCAATAAGTACATGACAGCAA 600
Db      |||
607 ACTGCTATGACAGCAGAGTATGAGTACGAGAAACAGAGAAATCAATAAGTATATGACAGCAA 666
Qy      |||
601 GGCAGAGTCCAGAGAGTATACATCTCCACCTGTATGAGAGTGCATTCATCACTCACA 660
Db      |||
667 GGTGAGTCCAAAGAAATACATCTCCACCTGTATGAGAGTGCATTCATCACTCACA 726
Qy      |||
661 ATCAGACATCTCTGGCAGTCCGCAATTTCTTAGTTAGCGAGCTTAAAGAGAGCGCGCAAT 720
Db      |||
727 ATTAGACATCTCTGGCAGTCCGCAATTTCTTAGTTAGCGAGCTTAAAGAGAGCGCGCAAT 786
Qy      |||
721 AGGAGAGTGGAGCTCCAGTATTAACAATCTAGTAGGGGATGTAGATCATCATCAGG 780
Db      |||
787 AGCAGAGTGGAGCTCCAGTATTAACAATCTAGTAGGGGATGTAGATCATCATCAGG 846
Qy      |||
781 AACCCGGACTTACTGCTATCTCTTCACTCAAAATATGGAATTAATACCAAGACATCA 840
Db      |||
847 AACCCGGCTTACTGCTATCTCTTCACTCAAAATATGGAATTAATACCAAGACATCA 906
Qy      |||
841 GGCCTAGCACTCAGCAGCCTCAAGCGGATATCCAAAGATCAAGCAGCTCATGCGTTTA 900
Db      |||
907 GGCCTGCACTCAGCAGCCTCAAGCGGATATCCAAAGATCAAGCAGCTCATGCGTTTA 966
Qy      |||
901 TATCGATGAAGGAGAAATGGCGGTATCATGCAATTCCTAGGTGACAGTATCAGATG 960
Db      |||
967 TATCGATGAAGGAGAAATGGCGGTATCATGCAATTCCTAGGTGACAGTATCAGATG 1026
Qy      |||
961 AGCTTTGCAACCGCTGAGTATGCACAGCTTTATCTTTTCCATGGCGATGGCATCAGTC 1020
Db      |||
1027 AGCTTTGCAACCGCTGAGTATGCACAGCTTTATCTTTTCCATGGCGATGGCATCAGTC 1086
Qy      |||
1021 TTAGATAAAGGAACCTGGCAATACCAATTCGCGAGAGCTTCATGAGCAGCATCATCTCG 1080
Db      |||
1087 TTAGATAAAGGAACCTGGCAATACCAATTCGCGAGAGCTTCATGAGCAGCATCATCTCG 1146
Qy      |||
1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGAGTATGATCAACGAGCAGCATGCTGCT 1140
Db      |||
1147 AGACTTGGAGTAGATATGCTCAGGCTCAGGAGCAGCATCAATGAGGACATGGCTGCT 1206
Qy      |||
1141 GAGCTAAACTAACCCCGCAGCAAGAGGGGCTGGCAGCTGCTGCCCAACGAGTGTCT 1200
Db      |||
1207 GAGCTAAACTAACCCCGCAGCAAGAGGGGCTGGCAGCTGCTGCCCAACGAGTGTCT 1266
Qy      |||
1201 GAGGAAACTGGCAGCTGGATATTCCTACTCAACAGCCGGGTCTCTCATCTGGGCTCAGC 1260
Db      |||
1267 GAGGAAACTGGCAGCTGGATATTCCTACTCAACAGCCGGGTCTCTCATCTGGGCTCAGC 1326
Qy      |||
1261 GATGAGGCCCCGAGCTCTCAGGCTGATCGAAACAAGTCGCAAGGGCAACAGATGCG 1320
Db      |||
1327 GACGGAGGCCCAAGCCCCCAAGCCGATCGAAACAAGTCGCAAGGGCAACCGGATGCG 1386
Qy      |||
1321 GGAGATGGGGAGACCCAAATCTTTGGATTTGATGAGAGCAGTGGCGCAACGAGCATGCGAGAA 1380
Db      |||
1387 GGGGATGGGGAGACCCAAATCTTTGGATTTGATGAGAGCAGTGGCGCAACGAGCATGCGAGAA 1446
Qy      |||
1381 GGCCAAACTCCGACAGAGACACACCCAGGGAAACCCCGGCTCTCCCGGGCCATCA 1440
Db      |||
1447 GGCCAAACTCCGACAGAGACACACCCAGGGAGCCATCCGAGGCTCCCCCAACTCTCGGGCCATCC 1506
Qy      |||
1441 CAGATAACGACACCGACTGGGGTATTGA 1470
Db      |||
1507 CCAGACACGACACTGACTGGGGTACTGA 1536
```

```
RESULT 13
AF431744 Newcastle Disease virus strain ZJ1, linear VRL 13-MAR-2003
LOCUS AF431744
DEFINITION AF431744
ACCESSION AF431744
VERSION AF431744.3 GI:28933797
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Huang, Y., Wu, Y., Wan, H. and Liu, X.
TITLE Complete nucleotide sequence of a NDV strain of goose origin (ZJ1)
JOURNAL Unpublished
AUTHORS Huang, Y., Wu, Y., Wan, H. and Liu, X.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-2001) Animal Infectious Disease Laboratory, Yangzhou University, 12 Wenhui East Road, Yangzhou, Jiangsu 225009, P.R. China
REFERENCE
AUTHORS Huang, Y., Wu, Y., Wan, H. and Liu, X.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-2003) Animal Infectious Disease Laboratory, Yangzhou University, 12 Wenhui East Road, Yangzhou, Jiangsu 225009, P.R. China
REMARK
AUTHORS Huang, Y., Wu, Y., Wan, H. and Liu, X.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2003) Animal Infectious Disease Laboratory, Yangzhou University, 12 Wenhui East Road, Yangzhou, Jiangsu 225009, P.R. China
REMARK
COMMENT Nucleotide and amino acid sequences updated by submitter
FEATURES
source
1. .15192
/organism="Newcastle disease virus"
/virion
/mol_type="genomic RNA"
/strain="ZJ1"
/db_xref="taxon:11176"
56..1798
/gene="NP"
56..1798
/gene="NP"
/product="nucleocapsid protein"
122..1591
/codon_start=1
/product="nucleocapsid protein"
/protein_id="AA18932.1"
/db_xref="GI:16417960"
/translation="MSSVFDEYEQLLAQTRPNHTGGGKSTLKVFPVFTLNSDD
EDGPNVSPQNNRSGVEERAFQFMVITAGSLPRACNSGTFVTAGVDEDPEDITD
PRDWNFVFCRLIAVSDANKPLRQGLISLCSHVSQVNRHVALAGQNEATLAVL
TIERILSQVQVMTVAKMTAYETADESETRINKYMQGQGVOKYIILHPVCRSAIQ
LTIRSLAVRIFLVBELKRRNTAGSSTYNNLGVDSYIRNTGLTAFELTKYGIN
TKTSALUSLLTGDIOKMKQLRMKMGENAPYMTLLGDSQMSFAPAEYIAQLYSFA
MGMAVLDKGTGKQFARDFMSGTFWRLGVEYAAQAGGSINEDMAELKLTAAIRGL
AAAQKRVSEIIGMDIPTQAGAGTGLSDGEPPTQGSNKPGQGPADGDTGTQFLDF
MEAVANSMEAPNPAQSTTHPEPPPTPGASQNDNDWGY"
1810..3250
/gene="P"
1810..3250
/gene="P"
/product="phosphoprotein"
1893..3080
/gene="P"
/codon_start=1
/product="phosphoprotein"
```



```
/protein_id="AAL18933.1"
/db_xref="GI:16417961"
/translaton="MAFTDAEIDDIFETSGTVIDSIIITAOQKSAETVERSAIPQKGT
KALSTAMEKHGSQPHASQADPQDQRTKQSTPBQVTLHNNPPISTGPPPTQAAS
ETSQTLTKGASNLSDMLDKLSNKSNAKSPMSQEGHHPQPAQHQHQPQSGYSGN
QGRPQHQAKAVPGRNGTDENTAYHQKESQLSAGATPHAPQSGQSDNTVPVDRVQ
LSADPAQAMMSMEALSQKSVKHOLDLVLKOTSSI PMMBSEIQOLKTSVAIMEANL
GMKILDPGCANVSSLSDLRAVASHPLVSGPDGPPSPYVYQGGEMTLNKLKSPVQHP
SELKSATASGPDGMGEKDTVRALITRPMFPSSAKLLSKLDAARSIEERIKRLA
LNG"
```

gene 3262. .4493

mRNA /gene="M"

3262. .4493

/gene="M"

/product="matrix protein"

3296. .4390

/gene="M"

/codon\_start=1

/product="matrix protein"

/protein\_id="AAL18934.2"

/db\_xref="GI:28911873"

```
/translaton="MDSRTIGLYFDSALPSSLLAFPIVLQDGTGKKQITPQYRIQ
RLDSWTQKEDSVITTYFIQIGNEATVGINDNPKHELLSSAMLCGLGSPNDGQ
LVELARACLTGVVTKKSATNTERIVSVQAPRVLOS CMVANRYSSVNAKVHAKP
EKIPGSGTLEYKNFVSLTVVPRDRDVYRIPTAVLVKVGSSLYNALAVTIDVDVDPKS
PLVKSLSKSDSGYANLFLHIGLMSVDYDKGKVIIDKIEKIRLNLVGLSDVLP
SVLVKARGATKLLAPSSSGTACYPANASQVAKILMSQTAHLRSVRVLIQAGTQ
RAVATADHEVTSTKIERHAIKYNPRK"
```

gene 4504. .6285

mRNA /gene="P"

4504. .6285

/gene="P"

/product="fusion protein"

4550. .6211

/gene="P"

/codon\_start=1

/product="fusion protein"

/protein\_id="AAL18935.3"

/db\_xref="GI:28933798"

```
/translaton="MGSKLSTRIIPAPMLITRIMLIIGCIRTPSSDLRPLAAAGIIV
TGDVNTSVSSQSGSIIVKLPNMRDKEACAPLEAVNRTLTLLTPIGDSIRKI
QGSVTSQGRQKRFI GAVISGVALGATAQIAAALIQANNAAILRLKESIAA
TNEAVHRTVDSQLSVAGVMQFVNDQFNNTARELDCIKITQCVGELNLYLTETL
TVFQPIQTSALTOITQALYNLAGNMDHLITLGIQNLSSLSGLITGPPIY
DSQTLQIQNLVSVGNLNNRATYILETISVTIKGTASALVAKVTVQVGSVIEED
TSCIESDLDYCTRI VTLPMSPGIYISCLSGNTSACMYKTEGALATTPYMALKGSVIA
NCKTICRCRDPGGIISQNTGEAVSLIDRHLNVLSDGITLRSGBFDATPYQKNI SI
LDSQVITGNLIDISTELGNVNNSISNALDLAESNKLKVNVRLTSTGTSALITIVILT
VISLVGALSGLACYLWYKQKQAQKTLMLGNNTLDQMRATTRA"
```

gene 6327. .8318

mRNA /gene="HN"

6327. .8318

/gene="HN"

/product="hemagglutinin-neuraminidase"

6418. .8133

/gene="HN"

/codon\_start=1

/product="hemagglutinin-neuraminidase"

/protein\_id="AAL18936.2"

/db\_xref="GI:28911875"

```
/translaton="MDRAVNRVLENEREAKNTRVLRVIAVLLVMVMTLATSAAAL
AYSGASTPHDLAGISTVISKEDKVTLSLSQDVIDKIKYQVLESPLALLNTESI
IMNATISLSQINGAANNSSCGAPVHPDPIYIGGKRLI VDDISDVTSPFSAYQEH
NFIPAPITGSGCQINGAANNSSCGAPVHPDPIYIGGKRLI VDDISDVTSPFSAYQEH
VFTSTLSTLNDTQNRKSVSATPLGCDMLCSKVTEEBEEDYKSVAPTSVMHGRG
FDGYHREKLDITTVLFDKVMVANYPGAGGFIIDRVWFPVYVGLKPNPSPSDPAQEGY
VIYKRHNTPCDEODYOIRMAKSSYKPRFGKQVQQAIIISIKTISLGDVLTIPP
NTLTLMAEGRILTVGTSHPFLYQRGSSYFSPALLYPMVNKNTATLSRPTFNAFTPP
GSVPQASACRPNCSITGYDDPYPLIFHRNHLRUGVFGTMDDEQALNPPSVAFVN
ISRSRTRVSSSSSKAAATYSTCFKVVTKNTKNTCLISAEISNTLFGFVPLVLEIL
KDDR"
```

gene 8376. .15079

mRNA /gene="L"

mRNA

8376. .15079

/gene="L"

/product="large polymerase protein"

8387. .15001

/gene="L"

/codon\_start=1

/product="large polymerase protein"

/protein\_id="AAL18937.2"

/db\_xref="GI:28911876"

```
/translaton="MAGSPERAHEHQIILPESHLSSPLVKHLLYVWKLTLGLPLPDEC
DEFHLIISROMKRILEGATPDMTRKIGRAVHOTLHNSKITGVHLPRCLLEELAS
VPDSTNKRERKLEKKIOIHNTYXGDLFTWKLCTHVEKKLIGSSRSNNRPSRSEPSIRTD
PAFWFKNSRAKFAWLHINQVQRHLI VAAIRAVNKLIVLSIKHGFVFPPELVIV
THTDENKFTCLTQELVLMYADMEDGDMVNIISSTAHLNLSKIDILRLDALAK
DLGNQYDVVALMEGFAYGQLLEPGTTFAGDFAFNLOELKDTLIELLPNNTAESV
THAATVFSGLEQNAEMLCLRLHGWPLLESRAFAVRSQCAPKMWDFMTLOV
LSFPKGTIINGYRKNKSGVPRVKVDTIYGNIIQLHADSAEISHDVMLRKYKSAL
EFEPICIDVPYNLSMPLKDKAIJHPSDNMLASPRNLSEDQKQIKEATSTRLLI
EFLSNPDFPKENEYILTLELRDDSVAVSYLKEKEVKNVGRIFAKLTKKLRNCQV
MABGILADQIAPFFQGVNGVODSISLTKMLAMSQLSFNKNKRIITDCKBRVSNRNH
DQSKNRRRVATFITDLOKYCLNRYQTVKLFAMHNLQMLPHFFEWIHLRMDTT
MFVGDPNPSPDPTDCDSLRVPNDIIVTSARGGIEGLCQKLTWMSIAAQLAARS
HCRVACMVQGNQVIAVTRVSRDDSDPMVLTQLHQAADNPFKELIHNHLIYHNLKD
RETRSDPTFYKRIFKDGAISLQVLKNSKLVLSGDSSENVTWMSCANASTVARL
CENGLPKDFCYLYNLMSCVQTYFDSFSTHSSQSOSNQSWIEDISFVHSYVLTPAQ
LGGUSNQYSLKTRNIGDPTTAPAEVKRLEAVGLLSPSITMNLITRPPGNDWASL
ENIDPNFETVASPNVLKKAQVLFETCSNPLSGVHTEDNEAEKALAEFLNQ
EVIPRVAHAIMEASSVGRRKIQGLVDTNTVTKIALTRPLGIRKRLMINTSSMH
AELSRDDI FLNRSNHPVLSNMCSLTADYARNRSPLTGGRKILGVSNPDTIELV
AMELSVSGCTCKDGDROFTPHLPSNLTEDTDSKNPMRPVPLGSGTQERRAS
EGETLSVSGCTCKDGDROFTPHLPSNLTEDTDSKNPMRPVPLGSGTQERRAS
LAKTAMSPHYKAALRASSVLIIMAYGNEVNTAALINARSCNISSEYLLSLPLPT
AGNLQHRDDGITTMTPTASLYRVSYPVHISNDSQRLTEEGVKRGNVYVQQIMLGL
LSLIESLFPMTTTRTYDEITLHLSKFSFCCIRVEPVAVPFVPELLGLAPELRMVTNKM
YDPSI SERDFARLDLAI FKSYLENLESYSTLEMLNLSISSGKLIGOSVVSYDEAS
IKNDIIVYDNTNRWISAEONS DVVLFYEAALVELLNLIDCAYQVYLRVGLNNTVLVY
NDLYKNMGPILLNIAATISHPIIHSRLNAGLINHDSHQDLADIFVEVSKLIVSC
TRVVSGLYAGNKYDILLPPSVLDNLEKMLQLISRLCCLTYVLFATREIPKIRGLS
AEEKCSILTYLLSDAVKPLRPEQSVSSIMPNIITPPANLYMSRKSNLNIRERDR
DETLSLFPQPELLELVYLFGRGTASSWSYKASHLLSVPEVRCARHNSLYLAESGSAI
FEETLPNKEDYLVYLFGRGTASSWSYKASHLLSVPEVRCARHNSLYLAESGSAI
MSLLELHPHETIYNTLFGNEMPPQHFQPTPTQFLNSVYRNIAQEVPCQGVYQ
EFCLPWRNEAESDTSQKAVGITYSVVYRSVSLHCDLEIPGSGNSQDQLQATL
SLIAMSVRGEGVVIKLYVAMYFFHLLMNFPTPCSTKGTTLNSGACRGMBCYLI
FVWGILGGPTFVHEVVRMAKTLQRHGTLKSDEITLTKFTSQRRVTDILSSPLP
KMRLLENDAAALIEAGQPVRFPCAESLVSITLMDTQTTQIASHIDTVIRSVYM
EABGDLADTVFLFTPNYLSLTKQCTKQCTQILRVTLGLRAKDTNKVGDVISIV
LRGAVSDEDILPLRTYLRKSTCPKYLKAVLGI TKLKEMFTDTSLLYLTRAQKQFYMT
IGNAAKGYYSNDS"
```

ORIGIN

Query Match 85.1%; Score 1250.8; DB 14; Length 15192;  
Best Local Similarity 90.7%; Pred. No. 0;  
Matches 1333; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

Qy	1	ATGCTTCCTCGATATCGATGATACGACGAGCTCTCTCGCTGCTGACATCGCCCCAATGGA	60
Db	122	ATGTCGTCTGTTTTTCGACGAATACGACGAGCTCTCTGCTGCTGACAGCCGCCCTAACGGA	181
Qy	61	GCTCAGCGAGGGGAGAGAGAGGAGGAGCACTTTTAAGAGTTCGAGGTCCAGGTCCCATGTTT	120
Db	182	ACTCATCGAGGGGAGAGAGGAGGAGCACTTTAAAGTTGAGGTCCCATGATTTTACCCTA	241
Qy	121	AACAGTCACGATCAGGAAGATAGATGGAATTTTTCGGGTATTTCTGTCCTTCGGATTCGT	180
Db	242	AACAGTCATGATCCGGAAGATAGATGGAATTTTTCGGGTATTTCTGTCCTTCGGATTCGT	301
Qy	181	AGCGAGAGCCCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGTGTCGCCAT	240
Db	302	AGCGAGGATCCCAACAAACCACTCAGGCAAGGTGCTCTTATATCCCTCTTATGTCGCCAT	361
Qy	241	TCTCACTGATGAGGACCATGTTGCCCTTTCGAGGAGAAACAGATGAGGCGCACACTGACT	300
Db	362	TCTCAGGTGATGAGAACCATGTTGCCCTTTCGAGGAGAAACAGATGAGGCGCACACTGCT	421





```
|||||
662 ACTGATATGACAGCAGATGAGTCAGAAACAGAGATATAATAGTATATGAGCAA 721
QY
601 GCGAGAGTCCAGAGAGATACATCTCCACCTGTATGAGAGAGTGCAATTCACATCACA 660
Db
722 GGTGAGTTACAGAGAGTACATCTTCTATCTGTATGAGAGAGTGCAATTCACATCACA 781
QY
661 ATCAGACATCTCTGCGAGTCCGATTTCTTCTAGTTAGGAGCTTAAGAGAGCGCGCAAT 720
Db
782 ATCAGACATCTCTGCGAGTCCGATTTCTTCTAGTTAGGAGCTTAAGAGAGCGCGCAAT 841
QY
721 AGCGAGGTGGAGCTCCACGATTAACAACTTAGTAGGGGATGTAGATCATATCATCAGG 780
Db
842 ACAGAGGTGGAGCTTACATATACAACTTAGTAGGGGATGTAGATCATATCATCAGG 901
QY
781 AACACGGAGTACTGTGATCTCTCTTCACTCAAAATATGGAATTAATACCAAGACATCA 840
Db
902 AACACGGGCTTACTGCTATTTCTCTTAACTCAATATGGAATTAATACCAAGACGTC 961
QY
841 GGCCTAGCACTCAGAGCGCTCAGAGCGGATATCCAAAGATGAAGCAGCTCATGCGTTTA 900
Db
962 GGCCTGCACTCAGAGCGCTCAGAGTGATATCCAAAGATGAAGCAGCTCATGCGTTTA 1021
QY
901 TATCGATGAAGGAGAGAAATGCGCGTACATGACATTTCTAGGTGACAGTGCATCAGATG 960
Db
1022 TATCGATGAAGGAGAGAAATGCGCGTACATGACATTTCTAGGTGACAGTGCATCAGATG 1081
QY
961 AGCTTTGACCGGCTGAGTATGACAGCTTTATCTTTTGGCATGGGATGGCATCAGTGC 1020
Db
1082 AGCTTTGACCGGCTGAGTATGACAGCTTTATCTTTTGGCATGGGATGGCATCAGTGC 1141
QY
1021 TTATAGTAAGGAGTGGCAAAATACCAATTCGCGAGAGTTCATGAGCACATCATCTCG 1080
Db
1142 TTATAGTAAGGAGTGGCAAAATACCAATTCGCGAGAGTTCATGAGCACATCATCTCG 1201
QY
1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGAGTAGCATCAACGAGACATGGTGTCT 1140
Db
1202 AGACTTGGAGTAGATATGCTCAGGCCAGGGAAGTAGCATTAATGAGGACATGGTGTCT 1261
QY
1141 GAGCTAAACTAACCCGAGCAGCAAGAGGGGCTGGAGCTGCTGCCCAAGAGTGTCT 1200
Db
1262 GAGCTAAACTAACCCGAGCAGCAAGAGGGGCTAGAGCTGCTGCCCAAGAGTGTCT 1321
QY
1201 GAGGAACTGGCAGCGTGGATATCTTACTCAACAGCGGGGTCTCTCACTGGGCTCAGC 1260
Db
1322 GAAGAATCGGAGATGACATTCCTCACTCAACAGCGGGGAGTCTCTCACGGGCTCAGT 1381
QY
1261 GATGAGGCCCCGAGCCTTCTAGGTTGATCGAAACAAGTCGCAAGGGCAACAGATGCC 1320
Db
1382 GACGAGGCCCCGAACTCCACAGGGCGGATCAACAAGCGCAAGGGCAACAGATGCC 1441
QY
1321 GGNATGGGAGACCAATCTTGGATTTGATGAGAGCAGTGGCGAACAGCATGCCAGAA 1380
Db
1442 GGGATGGGAGACCAATCTTGGATTTTATGAGAGCAGTGGCGAACAGCATGCCAGAA 1501
QY
1381 GCGCCAACTCCGACAGAGCACACCCACCCGAGCTCCCGGGCCATCA 1440
Db
1502 GCGCCAACTCCGACAGAGCACACCCATCCAGAACCTCTCCAACTCTGGGGCATCC 1561
QY
1441 CAAGATAACACACCGAGCTGGGGTATGTA 1470
Db
1562 CAAGATAACACACTGACTGGGGTACTGA 1591
```

## RESULT 15

```
AF419400
LOCUS
DEFINITION
  Newcastle disease virus chicken/Honduras/44813/00 nucleocapsid
  protein gene, complete cds.
ACCESSION
  AF419400
VERSION
  AF419400.1 GI:15811667
KEYWORDS
  Newcastle disease virus
```

```
ORGANISM
  Newcastle disease virus
  Viruses; ssRNA negative-strand viruses; Mononegavirales;
  Paramyxoviridae; Paramyxovirinae; Rubulavirus.
REFERENCE
  1 (bases 1 to 1694)
  Seal, B.S., Crawford, J.M., Sellers, H.S., Locke, D.P. and King, D.J.
  Nucleotide sequence analysis of the Newcastle disease virus
  nucleocapsid protein gene and phylogenetic relationships among the
  Paramyxoviridae
  Virus Res. 83 (1-2), 119-129 (2002)
JOURNAL
  MEDLINE
  21854562
  11864745
REFERENCE
  2 (bases 1 to 1694)
  Seal, B.S., Crawford, J.M., Sellers, H.S., Locke, D.P. and King, D.J.
  Direct Submission
  Submitted (14-SEP-2001) Agricultural Research Service, USDA,
  Southeast Poultry Research Laboratory, 934 College Station Road,
  Athens, GA 30605, USA
FEATURES
  Location/Qualifiers
  1..1694
    /organism="Newcastle disease virus"
    /mol_type="genomic RNA"
    /isolate="chicken/Honduras/44813/00"
    /db_xref="taxon:11176"
  15..1484
    /codon_start=1
    /product="nucleocapsid protein"
    /protein_id="AA09053.1"
    /db_xref="GI:15811668"
    /translation="MSSVFDEVEQLLAQTRPENGAGGEGKSTLKVEVPVFTLNSDD
    PEDRNWVFCURIAVSDANKPLRQALISLCSHVSQVRNHHVALAGKONATLAIL
    EIDGFANSTPQNVNVSBEERAPFWMISLPRACSNSTPFTVAGVDDPDTID
    TLERILSIQAQIWNVAKAMTAYETADESETIRNKYMQQGVQKYYILHDPVCRSAIQ
    LTRISLAVRIPLVSELKGRNTAGSSYTNLVGSDVSYIRNTGLTAFPLTKYGIN
    TKTSLAULSLTGDJQKMLRMRYMKGENAPYMTLLGSDQMSGPAPAYALYFPA
    MGMASVLDKGTQYQFAPDFMSTFWRLGVEVAQOGSSINEDMAELKLTDAARGL
    ARAQRVSEISGMDVPTQQAGVLGSDVGFPQAGGSHNSQGGPDAGDGETQFLDM
    MRVANSRMREVSALSTPLAPPOTPGSPDNDTDWGY"
```

## ORIGIN

```
Query Match      84.8%; Score 1246; DB 14; Length 1694;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 1 ATGCTCTTCGATTCGATGAATACGAGCAGCTCTCGCTGCTCAGACTCGCCCCAATGGA 60
Db 15 AUGTCGCTCTGTTTGAACGAATACGAGCAACTCTAGCTGCTCAGACCCGCCCTAATGGA 74
QY 61 GCTCAGGAGGGGAGAGAGAGGAGCACTTTAAGAGTTGAGTCCCGATTCCTCTT 120
Db 75 GCTCATGAGGAGGAGAGAGGAGGAGCACTTTGAAAGTTGAGTCCCGATTCCTCTT 134
QY 121 AACAGTGCAGATCCAGAAGATAGATGGAATTTTGGCGTATTCCTCTTCGGATTGCTGTT 180
Db 135 AATAGTGCAGATCCAGAAGACAGATGGAATTTTTCGGTATTCCTCTTCGGATTGCGGT 194
QY 181 AGCGAGAGCGCCAAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGTGCTCCCAT 240
Db 195 AGCGAGATGTCACCAACAAAGCACTCAGGCAAGGTGCTCTTATATCCCTCTGTGCTCCCAT 254
QY 241 TCTCAAGTGCAGGAGAACCATGTTGCCCTTGCAGGAAACAGAAATCAGGCTACAGTACT 300
Db 255 TCTCAAGTGCAGGAGAACCATGTTGCCCTTGCAGGAAACAGAAATCAGGCTACAGTACT 314
QY 301 GTTCTTTGAGATCGATGGTTTACAGCAGCGTGCCTCAGTTTCAACCAACAGAGAGTGGGGTG 360
Db 315 ATTCTTTGAGATTCATGGTTTGCACACAGCACCGCTCAGTTTCAACCAACAGAGAGTGG 374
QY 361 TCTGAGAGAGAGACACAGAGATTCATGCTGATGAGAGGCTCTCTCCCTCGGGGTGAGT 420
Db 375 TCTGAGAGAGAGAGACAGAGATTCATGCTGATGAGAGGCTCTCTCCCTCGGGGTGAGT 434
QY 421 AACGGTACTCCGTTCTGTCAGCGGTGGGTTGAAGATGATGCACACAGAAATATCAGTACT 480
```

Db 435 AACGGTACCCCGTTCTGTCACAGCCGGGTTGAAGATGATGCACACAGACATCACTGAC 494  
QY 481 ACTCTGGAAGAATCTGTCTATCTCAGGCTCAGGTATGGTTCACAGTAGCGAGGCCATG 540  
Db 495 ACTCTGGAAGAAGATCTGTCTATCTCAGGCTCAGATATGGTTCACAGTAGCGAGGCTATG 554  
QY 541 ACTGCATATGACACAGCAGATGAGTCGGAACAAGAGAATCAATAAGTACATGCAGCAA 600  
Db 555 ACTGCATATGAACAGCAGATGAGTCAGAAACAAGAGATCAATAAGTACATGCAGCAA 614  
QY 601 GGCAGATCCAGAGAAGTACATCTCCACCTGTATGCAGGATGCATCAACTCACTACA 660  
Db 615 GGCAGATCCAGAGAAGTACATCTCCACCTGTATGCAGGATGCATCAACTCACTACA 674  
QY 661 ATCAGACATTTCTCTGGCAGTCGCACTTTCTTAGTTAGGAGCTTAAAGAGGCGCGAAT 720  
Db 675 ATCAGGCAATCTTTGGCAGTTCCGCAATTTCTTAGTTAGGAGCTTAAAGAGGCGCGAAT 734  
QY 721 ACGGAGGTGGGAGCTCCAAGTATTAACAATTTAGTAGGGATGTAGACTCATACATCAGG 780  
Db 735 ACGGAGGTGGGAGCTCCACATATTACAACCTTGGTAGGGATGTAGACTCATACATCAGA 794  
QY 781 AACACCGCACTTACTGCAATTTCTCTTACATCTCAAAATATGGAATTAATCAAGACATCA 840  
Db 795 AACACTGGACTTACTGCAATTTCTCTGACACTTAAATATGGAATTAATCAAGACATCA 854  
QY 841 GCCCTAGCACTCAGCAGCTCACAGCGGATATCCAAAAGATCAAGCAGCTCATGCGTTTA 900  
Db 855 GCTCTTGCACTCAGCAGCTCACAGGTGATATCCAAAATTAAGAGCACTCATGCGTTTA 914  
QY 901 TATCGGATGAAGGAGGAAATGCGCGTACATGACATTTAGTGTGACAGTGTATCAGATG 960  
Db 915 TATCGGATGAAGGAGGAAATGCAACCTGACATGACACTGCTAGGTGACAGCGACAGATG 974  
QY 961 AGCTTTGACCGGTGAGTATGCAAGCTTTATTTTGGCATGGGCATGGCATFCAGTC 1020  
Db 975 AGCTTTGACCGGTGAGTATGCAACCTTTATTTTGGCATGGGCATGGCATFCAGTC 1034  
QY 1021 TTAGATAAGGAATGGCAATACCAATTCGCCAGAGACTTCATGAGCAGATCATCTTGG 1080  
Db 1035 TTAGATAAGGAATGGCAATACCAATTCGCCAGAGACTTCATGAGCAGATCATCTTGG 1094  
QY 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAAGACATGGCTGCT 1140  
Db 1095 AGGCTCGGGTGGAGTATGCTCAGGCTCAGGGAAGTAGCATTAATGAGGATATGGCTGCT 1154  
QY 1141 GAGCTAAACTTAACCCCGCAGCAAGAGGGGCTTGGCAGCTGCTGCCCAACGAGTGTCT 1200  
Db 1155 GAGCTAAAGCTTAACCCCGCAGCAAGGAGGGGCTTAGCAGCTGCTGCCCAACGAGTGTCT 1214  
QY 1201 GAGGAACTGGCAGCGTGGATATTCCTACTCAACAGCGGGTCTCCTCCTGGGCTCAGC 1260  
Db 1215 GAGGAACTGGCAGCGTGGATATTCCTACTCAACAGCGGGGAGTCTCCTCCTGGGCTCAGC 1274  
QY 1261 GATGGAGGCCCGGAGCTCTCAGGGTGGATTCGAACGAAGTGCAGAGGCAACCAAGATGCC 1320  
Db 1275 GACGTAGGCCCTCAAGCCCAAGCGGATCGCAATTCACAAGGCAACCGATGCC 1334  
QY 1321 GGAGATGGGAGACCAATTTCTTGATTTGATGAGAGAGTGGCGAACAGCATGCGAGAA 1380  
Db 1335 GGGAGCGGAGAGACCAATTTCTTGATATGATGAGGGCGGTGGCAATAGCATGCGAGAA 1394  
QY 1381 GCGCAAACTCCGACAGAGACCAACCCAGCAACCCCGGAGTCCCGGAGTCCCGGGCCATCA 1440  
Db 1395 GTACAAAATTTCTGACATGAGCACCACCCCTCTAGCGCTCTCCCAAACTCTCGGGCCATCC 1454  
QY 1441 CAAGATAACGACACCGACTGGGGTATTGA 1470  
Db 1455 CCAGACAATGACACCGACTGGGGTACTGA 1484

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 19:58:50 ; Search time 997 Seconds  
(without alignments)  
7328.827 Million cell updates/sec

Title: US-09-970-851-1  
Perfect score: 1470  
Sequence: 1 atgtctccgatttcgatga.....acaccgactgggggtattga 1470

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
  - 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
  - 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
  - 17: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
  - 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
  - 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1162.8	79.1	15882	13	US-10-440-419-56
2	1156.4	78.7	1801	13	US-10-415-981-3
3	1154.8	78.6	15900	13	US-10-440-419-55
4	1148.4	78.1	15186	10	US-09-741-744A-134
5	1148.4	78.1	15186	15	US-10-377-718-3
6	1148.4	78.1	15186	16	US-10-429-735-3
7	106.6	7.3	1572	10	US-09-951-061A-125
8	100.2	6.8	1578	10	US-09-951-061A-134
9	76	5.2	15456	10	US-09-900-112-35
10	76	5.2	15456	10	US-09-900-112-36
11	73.2	5.0	15609	17	US-10-302-547-131
12	72.2	4.9	15492	9	US-09-733-692A-60
13	72.2	4.9	15492	9	US-09-733-692A-62
14	72.2	4.9	15498	9	US-09-733-692A-61

15	39.8	2.7	11809	10	US-09-764-891-10172	Sequence 10172, A
16	39.2	2.7	299	15	US-10-184-644-164	Sequence 164, App
17	39.2	2.7	299	15	US-10-184-634-164	Sequence 164, App
18	39.2	2.7	14055	15	US-10-156-761-2883	Sequence 2883, Ap
19	39.2	2.7	125746	15	US-10-156-761-15102	Sequence 15102, A
20	39.2	2.7	9025608	15	US-10-156-761-1	Sequence 1, Appli
21	38.2	2.6	12584	13	US-10-221-613-361	Sequence 361, App
22	37.2	2.5	777	15	US-10-184-644-348	Sequence 348, App
23	37.2	2.5	777	15	US-10-184-634-348	Sequence 348, App
24	36.8	2.5	1484	13	US-10-424-599-59394	Sequence 59394, A
25	36.8	2.5	68750	14	US-10-014-717-1	Sequence 1, Appli
26	36.2	2.5	4934	17	US-10-475-970-2	Sequence 2, Appli
27	36	2.4	703	13	US-10-027-632-126919	Sequence 126919,
28	36	2.4	703	16	US-10-027-632-126919	Sequence 126919,
29	36	2.4	760	17	US-10-437-963-16184	Sequence 16184, A
30	36	2.4	923	13	US-10-027-632-120542	Sequence 120542,
31	36	2.4	923	13	US-10-027-632-120543	Sequence 120543,
32	36	2.4	923	13	US-10-027-632-120544	Sequence 120544,
33	36	2.4	923	16	US-10-027-632-120543	Sequence 120543,
34	36	2.4	923	16	US-10-027-632-120544	Sequence 120544,
35	36	2.4	923	16	US-10-027-632-120543	Sequence 120543,
36	35.8	2.4	524	15	US-10-029-386-11686	Sequence 11686, A
37	35.8	2.4	594	13	US-10-142-426-10	Sequence 10, Appl
38	35.8	2.4	594	15	US-10-123-155-10	Sequence 10, Appl
39	35.8	2.4	594	15	US-10-146-731-10	Sequence 10, Appl
40	35.8	2.4	594	15	US-10-140-472-10	Sequence 10, Appl
41	35.8	2.4	594	15	US-10-141-761-10	Sequence 10, Appl
42	35.8	2.4	594	15	US-10-142-885-10	Sequence 10, Appl
43	35.8	2.4	594	15	US-10-158-790-10	Sequence 10, Appl
44	35.8	2.4	594	16	US-10-137-871-10	Sequence 10, Appl
45	35.8	2.4	594	16	US-10-140-923-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-10-440-419-56  
; Sequence 56, Application US/10440419  
; Publication No. US20030224017A1  
; GENERAL INFORMATION:  
; APPLICANT: SAMAL, SIBA K.  
; APPLICANT: HUANG, ZHUHUI  
; TITLE OF INVENTION: RECOMBINANT NEWCASTLE DISEASE VIRUSES USEFUL AS  
; TITLE OF INVENTION: VACCINES OR VACCINE VECTORS  
; FILE REFERENCE: 108172-00096  
; CURRENT APPLICATION NUMBER: US/10/440,419  
; CURRENT FILING DATE: 2003-05-19  
; PRIOR APPLICATION NUMBER: 09/926,431  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: PCT/US00/06700  
; PRIOR FILING DATE: 2000-05-05  
; PRIOR APPLICATION NUMBER: 60/381,462  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/171,072  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: 60/132,597  
; PRIOR FILING DATE: 1999-05-05  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 15882  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant cDNA for NDV  
US-10-440-419-56

Query Match 79.1%; Score 1162.8; DB 13; Length 15882;  
Best Local Similarity 86.9%; Pred. No. 0;  
Matches 1278; Conservative 0; Mismatches 192; Indels 0; Gaps 0;  
QY 1 ATGCTTCGGTATTCGATGATACGAGACTCTCGTCTGCTCAGACTCGCCCAATGGA 60



Db 122 ATGCTTCCTCGATTTGACGAGTACGAACAGCTCTCTCGGCTCAGACTCGCCCCAATGGA 181  
QY 61 GCTCAGGAGGGGAGAGAGAGGAGCCTTTAAAGATTGAGGTCCCAAGTATTCACCTTT 120  
Db 182 GCTCATGAGGAGGGGAAAAGGGAGTACCTTTAAAGTAGACGTCCCGGATTCACCTTT 241  
QY 121 AACAGTGCATCCAGAGATGATGGAATTTTGGGTATTTCTGTCTTCGGATTCCTGTT 180  
Db 242 AACAGTGCATCCAGAGATGATGGAATTTTGGGTATTTCTGTCTTCGGATTCCTGTT 301  
QY 181 AGCAGAGCGCCAAACAAACCGCTCAGGCAAGGTCTCTCATATCCCTCTGTGCTCCCAT 240  
Db 302 AGCAGAGTGCACAAACCACTCAGGCAAGGTCTCTCATATCTCTTTATGCTCCAC 361  
QY 241 TCTCAAGTGCATGAGGAACCATGTCCTTTCAGGAAACAGAAATGAGGCTACATGACT 300  
Db 362 TCACAAGTGCATGAGGAACCATGTTGCCCTTCAGGGAACAGAAATGAGGCAATGGCC 421  
QY 301 GTTCTTGAGATGATGGTATTTACAGCAGCGTCCCTCAGTTCAAACAAGGAGTGGGTG 360  
Db 422 GTGCTTGAGATGATGGCTTTGCCAACGGTATGCCCCAGTTCAAACAATGAGGATGAGTG 481  
QY 361 TCTGAGGAGAGCAGAGATTCATGGGTAGCAGAGGTCTCTCCTCGGCGTGCAGT 420  
Db 482 TCTGAGGAGAGCAGAGATTCGCGATGATGAGGCTCTCTCCTCGGCGTGCAGT 541  
QY 421 AACGCTACTCCGTTCTGTCAGGCTGGGTGGAAGATGATGCAACCAAGATATCACTGAT 480  
Db 542 AATGGCACCCCTTCTGTCACAGCGGGGCGAAGATGATGCAACCAAGATATCACTGAT 601  
QY 481 ACTCTGGAAGATCTCTGTCTATCCAGGCTCAGGTATGGTCAAGTATGGTCAAGTATGGT 540  
Db 602 ACCCTGAGAGATCTCTCTATCCAGGCGCCAGTATGGTCAAGTATGGTCAAGTATGGT 661  
QY 541 ACTGTCATATGAGACAGATGATGTCGGAACCAAGAAATCAATAGTACATGACGAA 600  
Db 662 ACTGCTATGAGCTGACATGATGCTGAAACAAGACGATCAGTAAATATGACGAA 721  
QY 601 GGCAGAGTCCAGAAAGTACATCTCTCCACCTGTATGAGGAGTGCATTCACCTCA 660  
Db 722 GGCAGGCTCCAAAGAAATACATCTCTACCCCGTATGAGGAGCAATCCAACTCAG 781  
QY 661 ATCAGACATCTCTGAGCTCCGATTTCTAGTTAGGAGCTTAAGAGAGCCGCAAT 720  
Db 782 ATCAGACATCTCTGAGCTCCGATTTCTAGTTAGGAGCTTAAGAGAGCCGCAAT 841  
QY 721 ACGGCAAGTGGAGCTCCAGTATTAACCTTAGTAGGGGATGTAGACTCATACATCAG 780  
Db 842 ACGGCAAGTGGTACTCTACTTATATACCTTAGTAGGGGACGTAGACTCATATATCAG 901  
QY 781 AACACCGGACTTACTGATTTCTCTTAACCTCAATATGGAATTAATACCAAGATCA 840  
Db 902 AATACCGGGCTTACTGCAATTTCTTTGACACTCAAGTACGGAATTAACCAAGATCA 961  
QY 841 GGCCTAGCACTAGCAGCTCAGAGGATATCCAAAGATGAGGAGCTCATGCGTTTA 900  
Db 962 GGCCTTGCATTTAGTACCTCTCAGGCGCATCCAGAAATGAGGAGCTCATGCGTTTA 1021  
QY 901 TATCGGATGAAGGAGAAATGGCCGTATCATGACATTTGCTAGGTGACAGTATCAGATG 960  
Db 1022 TATCGGATGAAGGAGATATGGCCGTATCATGACATTTGCTAGGTGATGACAGATG 1081  
QY 961 AGCTTTGACCGGCTGAGTATGACAGCTTATTTCTTTGCAATGGCATGGCATCAGTC 1020  
Db 1082 AGCTTTGCGCTGCGGATGACACAACTTTACTCTCTTCGCCATGGGTATGGCATCAGTC 1141  
QY 1021 TTAGATAAGGAATGCAATACCAATTCGAGAGACTTCATGACACATCATTTCTGG 1080  
Db 1142 CTAGATAAGGATCTGGGAATACCAATTTGCGAGGACTTTATGAGCAATCATTTCTGG 1201  
QY 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGATGAGCATCAACGAAGCATGGCTGCT 1140

Db 1202 AGACTTGGAGTAGAGTACGCTCAGGCTCAGGGAAGTAGCAATTAACGAGGATATGGTGCC 1261  
QY 1141 GAGCTAAACTAAACCCCGGAGCAGAAAGGGGCTCGGAGCTGCTGCCAACGAGTGTCT 1200  
Db 1262 GAGCTAAAGTTAAACCCCGGAGCAGAAAGGAGGCGCTGGCAGCTGTCTGCCAACGAGTGTCT 1321  
QY 1201 GAGGAAACTGGCAGCGTGTGATTTCTTACTCAACAAGCCGGGCTCTCCTCAGTGGGCTCAGC 1260  
Db 1322 GAGGAGACACGAGCATAGACATGCTACTCAACAAGTCGGAGTCTCTCAGTGGGCTCAGC 1381  
QY 1261 GATGAGGCCCCCGGAGCCTCTCAGGCTGATGGAACAAGTCGCAAGGGGCAACAGATGCG 1320  
Db 1382 GAGGGGGGCTCCCAAGCCCTACAGGCGGATCGAATAGATCGCAAGGGGCAACAGAAAGCC 1441  
QY 1321 GGAGATGGGAGACCCCAATTTCTTGGATTTGATGAGGAGTGGCGAAACAGCATGCGAGAA 1380  
Db 1442 GGGGATGGGAGACCCCAATTTCTTGGATTTGATGAGGAGTGGCGAAATAGCATGAGGAA 1501  
QY 1381 GGGCCAAACTCCGCAACAGACACCCACCCCGGAAACCCCGGCACTCCCGGGCCATCA 1440  
Db 1502 GGGCCAAACTCTGCACAGGCGACTCCCAATCGGGGCTCTCCCGCAACTCTCGGGCATCT 1561  
QY 1441 CAAAGTAACGACACCGACTGGGGGTATTGA 1470  
Db 1562 CAAAGTAACGACACCGACTGGGGGTATTGA 1591

## RESULT 2

US-10-415-981-3  
; Sequence 3, Application US/10415981  
; Publication No. US20040043035A1  
; GENERAL INFORMATION:  
; APPLICANT: AKZO No. US20040043035A1el N.V.  
; TITLE OF INVENTION: A recombinant Newcastle disease virus nucleoprotein  
; FILE REFERENCE: 2000583  
; CURRENT APPLICATION NUMBER: US/10/415,981  
; CURRENT FILING DATE: 2003-05-02  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1801  
; TYPE: DNA  
; ORGANISM: Newcastle disease virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (122)..(1588)  
; OTHER INFORMATION: NP gene: nucleotides 56-1801; NP coding sequence:  
; OTHER INFORMATION: nucleotides 122-1588  
US-10-415-981-3

Query Match 78.7%; Score 1156.4; DB 13; Length 1801;  
Best Local Similarity 86.7%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 196; Indels 0; Gaps 0;  
QY 1 ATGCTTCCTCGATTTTCGATGAATACGAGCAGCTCCTCGCTGCTCAGACTCGCCCCAATGGA 60  
Db 122 ATGCTTCCTCGATTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181  
QY 61 GCTCAGGAGGGGAGAGAGGAGGAGCCTTTAAGAGTTGAGGTCCAGTATTCACCTTT 120  
Db 182 GCTCATGAGGGGAGAGAGGAGGAGTACCTTTAAAGTAGTAGCTCCCGGTATTCACCTTT 241  
QY 121 AACAGTGCATCCAGAGATGATGGAATTTTGGGTATTTCTGTCTTCGGATTCCTGTT 180  
Db 242 AACAGTGCATCCAGAGATGATGGAATTTTGGGTATTTCTGTCTTCGGATTCCTGTT 301  
QY 181 AGCAGAGCGCCAAACAAACCGCTCAGGCAAGGTCTCTCATATCCCTCTGTGCTCCCAT 240  
Db 302 AGCAGAGTGCACAAACCACTCAGGCAAGGTCTCTCATATCTCTTTATGCTCCAC 361  
QY 241 TCTCAAGTGCATGAGGAACCATGTTGCCCTTGGAGGAAACAGAAATGAGGCTACACTGCT 300



Db 362 TCACAGGTAATGAGGAACCATGTTGCCCTTGCAGGGAAACAGATGAGCCACATTTGGCC 421  
Qy 301 GTTCTTTGAGATCGATGGTTTACACGACGCGTCCCTCAGTTTCAACAAAGGAGTGGGTG 360  
Db 422 GTGCTTTGAGATGATGGCTTTTCCCAACGGCAACGCCCTCAGTTTCAACAAATAGGATGGAGTG 481  
Qy 361 TCTGAGGAGAGACACAGAGATTCATGTGTATAGCAGGGGTCTCTCCCTCGGCGTGCAGT 420  
Db 482 TCTGAAGAGAGAGACACAGAGATTTGCGATGATAGCAGGATCTCTCCCTCGGCGATGCAGC 541  
Qy 421 AACGGTACTCGTTTGTTCAGCGCTGGGGTGTGAAGATGATGACACAGAGATATCACTGAT 480  
Db 542 AACGGAAACCCGGTTTGTTCACAGCCGGGGCCGAAGATGATGACACAGAGATATCACTCCGAT 601  
Qy 481 ACTCTGAAAGAACTCTGTCTATCAGGCTCAGGTATGGGTACAGTACGAGGCCATG 540  
Db 602 ACCCTGGAGAGATCTCTCTATCAGGCTCAAGTATGGGTACAGTACGAGGCCATG 661  
Qy 541 ACTCGATATGACAGCAGATGAGTTCGGAACCAAGAGAAATCAATAGTACATGACGAA 600  
Db 662 ACTCGTATGAGACTGCAGATGAGTTCGGAACCAAGAGAAATCAATAGTATATGACGAA 721  
Qy 601 GGCAGTCCAGAGAGATGATCTCTCAGCCCTGTATGAGAGTGCATCACTCA 660  
Db 722 GGCAGGTTCAGAGAAATATATCTCTACCCCGTATGACAGGACCAATCACTCAG 781  
Qy 661 ATCAGACATCTCTGGCAGTCCGATTTTCTAGTTAGGAGCTTAAGAGGGCCGCAAT 720  
Db 782 ATCAGACATCTCTGGAGTCCGATCTTTTGGTTAGGAGCTCAAGAGGGCCGCAAC 841  
Qy 721 ACGCAGGTGGAGTCCAGCTATTACAACTTAGTAGGGGTGTAGACTACATACAGG 780  
Db 842 ACGCAGGTGGTACTCTTACTTATTATTAACCTGTAGGGAGCTAGACTACATCAGG 901  
Qy 781 AACACCGGACTTACTGATCTTCTTACATCTCAATATGAAATTAATACCAAGATCA 840  
Db 902 AATACCGGGCTTACTGATCTTCTTGACACTCAAGTACGGAATCAACACCAAGACATCA 961  
Qy 841 GCCCTAGCACTCAGCAGCTCAGAGGGATATCAAAAGATGAGAGCTCATGGTTTA 900  
Db 962 GCCCTTGCATTTAGTACCTCTAGCGGACATCAGAAAGATGAAAGCACTCATGGTTG 1021  
Qy 901 TATCGGATGAGGAGGAGAAATCGCCGTACATGATGATTCGTAGCTGAGTGATCAGATG 960  
Db 1022 TATCGGATGAGGAGGATATGCGCGGTACATGATGATTTACTTGTGTAGTATGACAGATG 1081  
Qy 961 AGCTTTGCACGGCTGAGTATGCAAGCTTTATTTTGGCCATGGGATGGCATCAGTC 1020  
Db 1082 AGCTTTGCGCTGCGGATGATGCAAACTTTTACTCTTTGGCATGGGTATGGCATCAGTC 1141  
Qy 1021 TTAGATAAGGAACTGGCAATATCAATTCGCCAGAGACTTTCATGAGCATCATCTTGG 1080  
Db 1142 CTAGATAAAGTACTGGGAATATCAATTTGCCAGGACCTTTATGAGCATCATCTTGG 1201  
Qy 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGGATGAGCATCAACGAGACATGGCTGCT 1140  
Db 1202 AGACTTGGATGAGTATGCTCAGGCTCAGGAGTATGATTAACGAGGATATGGCTGCC 1261  
Qy 1141 GAGCTAAATCAACCCCGGAGAGAGAGGGGCTGCGAGTGTGTCCTCCACAGAGTGTCT 1200  
Db 1262 GAGCTAAAGTAAACCCCGAGAGAGAGGGGCTGCGAGTGTGTCCTCCACAGGGTCTCC 1321  
Qy 1201 GAGGAACTGCGAGTGGATATTTCTTACTCAACAGCGGGTCTCTCAGTGGCTCAGC 1260  
Db 1322 GAGGAGACGAGCATAGACATGCCCTACTCAACAGTGGAGTCTCTCAGTGGGCTTAGC 1381  
Qy 1261 GATGGAGGCCCGGAGCTCTCAGGGTGGATTCGAAACAAAGTCGAAAGGCGCAACAGATGCC 1320  
Db 1382 GAGGGGGGTCCCAAGCTCTACAGGCGGATCGATATAGTGCAGAGGCGCAACAGAGGCC 1441  
Qy 1321 GAGATGGGAGAGACCCAAATTTCTTGGATTTGATGAGAGCTGGCGAAACAGCATGCGAGAA 1380  
Db 1442 GGGGATGGGAGAGACCCAAATTTCTGGATCTGATGAGAGCGGTAGCAAAATAGCATGAGGGAG 1501

Qy 1381 GCGCAAACTCCGCAACAGAGACACCAACCCGGAACCCCGGCACTCCCGGGGCATCA 1440  
Db 1502 GCGCAAACTCTGCAACAGGGCACTCCCAATCGGGGCTCCCGCACTCTCGGGCATCC 1561  
Qy 1441 CAAGATAACGACACCGACTGGGGTATTGA 1470  
Db 1562 CAAGATAACGACACCGACTGGGGTATTGA 1591

## RESULT 3

US-10-440-419-55  
; Sequence 55, Application US/10440419  
; Publication No. US20030224017A1  
; GENERAL INFORMATION:  
; APPLICANT: SAMAL, SIBA K.  
; APPLICANT: HUANG, ZHUHUI  
; TITLE OF INVENTION: RECOMBINANT NEWCASTLE DISEASE VIRUSES USEFUL AS  
; TITLE OF INVENTION: VACCINES OR VACCINE VECTORS  
; FILE REFERENCE: 108172-00096  
; CURRENT APPLICATION NUMBER: US/10/440,419  
; PRIOR FILING DATE: 2003-05-19  
; PRIOR APPLICATION NUMBER: 09/926,431  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: PCT/US00/06700  
; PRIOR FILING DATE: 2000-05-05  
; PRIOR APPLICATION NUMBER: 60/381,462  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/171,072  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: 60/132,597  
; PRIOR FILING DATE: 1999-05-05  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 15900  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant cDNA for NDV  
US-10-440-419-55

Query Match 78.6%; Score 1154.8; DB 13; Length 15900;  
Best Local Similarity 86.6%; Pred. No. 0;  
Matches 1273; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

Qy 1 ATGCTTCCGTTATTCGATGAATACGACAGCTCCCTCGCTGCTCAGACTCGCCCCCAATGGA 60  
Db 836 ATGCTTCCGTTATTCGATGAATACGACAGCTCCCTCGGCTCAGACTCGCCCCCAATGGA 895  
Qy 61 GCTCAGCGAGGGGAGAGAGGAGGAGCACTTTAAGAGTTGAGTCCCAAGTATTCACCTCTT 120  
Db 896 GCTCATGAGGGGAGAGAAAGGAGTACCTTAAAGTAGACGTCCTCGGATTCACCTCTT 955  
Qy 121 AACAGTACACCTCCAGAGATAGATGGAATTTTGGGTATTTCTCTCGGATTCGCTGTT 180  
Db 956 AACAGTATACCCAGAGATAGATGGAAGCTTTTGGTATTTCTGCTCCGATTCGCTGTT 1015  
Qy 181 AGCGAGGACGCCCAACAAACCGCTCAGGCAAGTGTCTCATATCCCTCTGTCCTCCAT 240  
Db 1016 AGCGAGATGCCCAACAAACCGCTCAGGCAAGTGTCTCATATCTCTTTTATGCTCCAC 1075  
Qy 241 TCTCAAGTATGAGGAACCAATGTTGCCCTTTCAGGAGAAACAGAAATGAGGCTACACTGACT 300  
Db 1076 TCACAGTATGAGGAACCAATGTTGCCCTTTCAGGAGAAACAGAAATGAGGCTACACTGACT 1135  
Qy 301 GTTCTTGAGATCCATGGTTTACAGAGGCTGCTCAGTTTCAACAGAGGTTGGGTG 360  
Db 1136 GTGCTTGAGATCCATGGTTTTCAGAGGCTGCTCAGTTTCAACAGAGGTTGGGTG 1195  
Qy 361 TCTGAGGAGAGACACAGAGATTCATGTTGATAGAGGGTCTCTCCCTCGGGGTGCGAT 420  
Db 1196 TCTGAGGAGAGACACAGAGATTCATGTTGATAGAGGGTCTCTCCCTCGGGGTGCGAT 1255

```
QY 421 AACGGTACTCTCGTTCGTCAAGCGCTGGGGTGAAGATGATGACACGAAGATATCACTCAT 480
Db 1256 AACGGAAACCCCGTTTCGTACACAGCCGGGCGGAAGATGATGACACGAAGATATCACTCAC 1315
QY 481 ACTCTGGAAGAAATCCTCTCTATCCAGGCTCAGGTATGGGTACACAGTAGCGAAGCCCATG 540
Db 1316 ACCCTGGAGAGATCCTCTCTATCCAGGCTCAAGTATGGGTACAGTAGCGAAGCCCATG 1375
QY 541 ACTGCATATGACACAGCAGATGAGTCGGAACCAAGAGAATCAATAAGTACATGCGACAA 600
Db 1376 ACTGGTATGAGCTGAGTAGCTGAGTCGGAACCAAGCGAATCAATAAGTATATGCGACAA 1435
QY 601 GGCAGAGTCGGAAGAGTATCTCTCAACCTGTATGACAGAGTGCAATTCATCTACA 660
Db 1436 GGCAGGGTCCAAAGAAATACATCTCTACCCCGTATGACGAGACCAATCCAACTCAG 1495
QY 661 ATCAGACATCTCTCGAGTCGCGATTTTCTTAGTTAGCGAGCTTAAGAGAGCGCGCAAT 720
Db 1496 ATCAGACAGTCTCTGAGTCGCGATCTTTTGGTTAGCGAGCTCAAGAGAGCGCGCAAC 1555
QY 721 ACGCAGGTGGAGCTTCAAGTATTAACAATTAGTAGGGATGTAGACTCATATCATCAGG 780
Db 1556 ACGCAGGTGGTACCTCTACTTATTATTAACCTGGTAGGGACGTAGACTCATATCATCAGG 1615
QY 781 AACACGGGACTTATCTGATCTCTCTTACATCAATATGGAATTAATACCAAGACATCA 840
Db 1616 AATACCGGGCTTACTGCAATCTCTTTGACACTCAAGTACGGAATCAACACCAAGACATCA 1675
QY 841 GCCCTAGCACTCAGACAGCTCAGCGGATATCCAAAGATCAAGCAGCTCATGGTTTGA 900
Db 1676 GCCCTTGCATTTAGTACCTCTCAGCGGACATCCAGAGTGAAGCAGCTCATGGTTTGA 1735
QY 901 TATCGGATGAAGGAGAAATGCGCGTATCATGACATTCCTAGGTGACAGTGAATCAGATG 960
Db 1736 TATCGGATGAAGGAGATAATGCGCGTATCATGACATTCCTAGGTGACAGTGAATCAGATG 1795
QY 961 AGCTTTGCAACCGGCTGAGTATGACAGCTTATCTTTGGCATGGCATGGCATGCAATCAGTC 1020
Db 1796 AGCTTTGCGCCCTGCGAGTATGACAACTTTACTCCCTTGCCATGGGTATGGCATCAGTC 1855
QY 1021 TTAGATAAAGGAATCGCAAAATACCAATTCGCGAGAGACTTCATAGACATCATCTCTGG 1080
Db 1856 CTAGATAAAGGTACTGGGAAATACCAATTTGCCAGGACTTTATAGACATCATCTCTGG 1915
QY 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGAGTAGCATCAAGAGACATGCTGCT 1140
Db 1916 AGACTTGGAGTAGAGTACGCTCAGGCTCAGGGAAGTAGCATTAACGAGGATATGGCTGCC 1975
QY 1141 GAGCTAAAGTAAACCCCGGACGAGAGAGGGGCTGCGAGCTGCTGCCAAACGAGTGTCT 1200
Db 1976 GAGCTAAAGTAAACCCCGGACGAGAGAGGGGCTGCGAGCTGCTGCCAAACGAGTGTCT 2035
QY 1201 GAGGAACTGGCAGCGTGGATATTCCTACTCAACAGCGGGGTCTCTCACTGGGCTCAGC 1260
Db 2036 GAGGAGACGAGCAGATAGACATGCTCTCAACAGTCGAGTCTCTCACTGGGCTTAGC 2095
QY 1261 GATGGAGGCCCCGAGCCTCTCAGGCTGAGTGAATGAAACAGTCGCAAGGGGCAACAGATGCC 1320
Db 2096 GAGGGGGGTCCCAAGCTCTCAAGGCGGATCGAATAGATCGCAAGGGGCAACAGAGGCC 2155
QY 1321 GGAGATGGGGAGACCAATTTCTTGGATTTTATGATGAGAGCAGTGGCGAACAGCATGCGAGAA 1380
Db 2156 GGGGATGGGGAGACCCAAATTTCTTGGATTTTATGATGAGAGCAGTGAATAGCATGAGGGAG 2215
QY 1381 GCGCCAACTCCGACAGACACCAACCCCGGAAACCCCGGACTCCCGGCGCATCA 1440
Db 2216 GCGCCAACTCTGACAGGCACTCCCGAATCGGGGCTCCCGCAACTCTCTGGGCGCATCC 2275
QY 1441 CAAGATTAACGACACCGGCTGGGGTATGA 1470
Db 2276 CAAGATTAACGACACCGGCTGGGGTATTTGA 2305
```

## RESULT 4

```
US-09-741-744A-134
; Sequence 134, Application US/09741744A
; Publication No. US20030087417A1
; GENERAL INFORMATION:
; APPLICANT: Peeters, Bernadus
; APPLICANT: de Leeuw, Olav
; APPLICANT: Klaus, Guus
; APPLICANT: Arnaud, Gielkens
; TITLE OF INVENTION: Newcastle Disease Virus Infectious Clones, Vaccines and Diagnost
; FILE REFERENCE: 2183-4646US
; CURRENT APPLICATION NUMBER: US/09/741,744A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PCT/NL99/00377
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 15186
; TYPE: DNA
; ORGANISM: Newcastle disease virus LaSota
US-09-741-744A-134
```

```
Query Match 78.1%; Score 1148.4; DB 10; Length 15186;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 201; Indels 0; Gaps 0;
```

```
QY 1 ATGCTTTCGGTATTTCGATGATATACGAGCAGCTCTCGCTGCTCAGACTCGCCCCAATGGA 60
Db 122 ATGCTTTCGGTATTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 181
QY 61 GCTCAGGAGGGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 182 GCTCATGAGGGGGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 241
QY 121 AACAGTGAAGTATCCAGAGATAGATGGAATTTTTCGGGTATTCGTCTTCGGATTCGTCTT 180
Db 242 AACAGTGAAGTATCCAGAGATAGATGGAATTTTTCGGGTATTCGTCTTCGGATTCGTCTT 301
QY 181 ACGGAGGAGCGCAACAAACCGCTCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
Db 302 ACGGAGAGTGCACAAACAAACCTCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 361
QY 241 TCTCAAGTATCAGGAGAACCATGTTGCCCTTCAGGAGAAACAGAAATGAGGCTACACTGACT 300
Db 362 TCACAGGTAATGAGGAGAACCATGTTGCCATTCAGGAGAAACAGAAATGAGGAGGAGGAGGAG 421
QY 301 GTTCTTGAGATCGATGGTTTACACGACGCTGCTCAGTTCACAAACAGGAGTGGGGTG 360
Db 422 GTGCTTGAGATTGATGGCTTTGCCAACGCGACGCCCCAGTTCAACAATAGGAGTGGAGTG 481
QY 361 TCTGAGGAGAGAGCAGAGATTCATGTTGATGAGAGGGTCTCTCCCTCGGGGTCAGT 420
Db 482 TCTGAGAGAGAGCAGAGATTTGCGATGATGAGAGGATCTCTCCCTCGGGATCGAGC 541
QY 421 AACGGTACTCTCGTTCGTCAAGGCTGGGGTTTGAAGATGATGACCAAGAGATATCACTGAT 480
Db 542 AACGGAAACCCCGTTCTGTCACAGCGGGGAGAGATGATGATGATGATGATGATGATGATGAT 601
QY 481 ACTCTGGAAGAAATCCTCTCTATTCAGGCTCAGGTATGGGTACAGTAGGAGGCGCATG 540
Db 602 ACCCTGGAGAGGATCTCTCTATTCAGGCTCAAGTATGGGTACAGTAGCAAAAGCGCATG 661
QY 541 ACTGCATATGAGACACAGATGATGCGGAAACAAAGAAAGAAATCAATAGATGATGATGATG 600
Db 662 ACTGCGTATGAGACTGCGAGATGATGCGGAAACAAAGCGGATCAATAGTATATGACACAA 721
QY 601 GGCAGAGTCCAGAGAGAGTACATCTCTCCACCTCTGATGACAGGAGTGAATTCATCACTCA 660
Db 722 GGCAGGGTCCAAAGAAATACATCTCTACCCCGTATGACAGGAGCAAAATCCCACTCAG 781
QY 661 ATCAGACATCTCTGCGAGTCCGCAATTTTCTTAGTAGGAGCTTAAGAGAGGCGCGCAAT 720
```

Db 782 ATCAGACAGTCTCTTGGAGTCCGACATCTTTTGGTTAGCGAGTCAAGAGAGCGCCGAC 841  
Qy 721 ACGGAGGTGGGAGCTCCAGGTATTACACTTTAGTAGGGAGTAGACTCATACATCAGG 780  
Db 842 ACGGAGGTGGGAGTCTTACTTATTATTAACCTGGTAGGGAGTAGACTCATACATCAGG 901  
Qy 781 AACACCGGACTTACTGCACTTCTTCTTACACTCAAAATATGGAATTAATACCAAGACATCA 840  
Db 902 AATACCGGCTTACTGCACTTCTTCTGACACTCAAGTACGGAATCAACCAAGACATCA 961  
Qy 841 GCCCTAGCACTCAGCAGCTCTCAGCGATATCCAAAGATGAAGCAGCTCATGCGTTTA 900  
Db 962 GCCCTTGCACCTTAGTAGCTCTCAGCGACATCCAGAGATGAAGCAGCTCATGCGTTG 1021  
Qy 901 TATCGGATGAAGGAGAAATCGCGGTACATGACATTTGTTGCCATGGGATGGCATCAGATG 960  
Db 1022 TATCGGATGAAGGAGAAATCGCGGTACATGACATTTGTTGCCATGGGATGGCATCAGATG 1081  
Qy 961 AGCTTTGACCGGCTGAGTATGCACAGCTTTTATCTTTTGCCATGGGATGGCATCAGTC 1020  
Db 1082 AGCTTTGCGCTCGCGAGTATGCACACTTTTACTCTTTTGCCATGGGATGGCATCAGTC 1141  
Qy 1021 TTAGTAAGGAATCGGCAATACCAATTCGCGAGACTTCATCAGCACAATCATTTGG 1080  
Db 1142 CTAGATAAGGATCTGGGAAATACCAATTTGCGAGGACTTTATGAGCACATCATTTCTG 1201  
Qy 1081 AGACTCGGGTGGAGTATGCTCAGGCTCAGGGAGTAGCATCAACGAAGACATGGCTGCT 1140  
Db 1202 AGACTTGGAGTAGAGTAGCTCAGGCTCAGGGAAGTAGCATTTAAGAGGATATGGCTGCC 1261  
Qy 1141 GAGCTAAACTAAACCCCGGACGAGAAAGGCGCTCGGAGCTGCTGCCAACGAGTGTCT 1200  
Db 1262 GAGCTAAAGCTAAACCCAGCAGCAATGAAGGCGCTGGCAGTGTCTGCCAACGAGTGTCT 1321  
Qy 1201 GAGGAATCGCAGGTGGATATTCCTACTCAACGAAGCGGGTCTCTCACTGGGCTCAGC 1260  
Db 1322 GACGATACGACAGCATATACATGCTACTCAACGAAGTGGAGTCTCTCACTGGGCTTAGC 1381  
Qy 1261 GATGAGGCGCCGAGCTCTCAGGTGGATGCAAGTGGAGTGGCAAGGCAACCGATGCC 1320  
Db 1382 GAGGGGGTCCCAAGCTCTCAAGGGGATGCGAATAGATGCAAGGGCAACCGAAGGCC 1441  
Qy 1321 GGAGATGGGAGACCCCAATCTTGGATTGATGAGAGCAGTGGCGAAACAGCATCGAGAA 1380  
Db 1442 GGGATGGGAGACCCCAATCTTGGATCTGATGAGGCGTAGCAATAGCATGAGGAG 1501  
Qy 1381 GCGCCAACTCCGACAGACCAACCCACCGGAAACCCCGGACTCCCGGGCCATCA 1440  
Db 1502 GCGCCAACTCTGACAGGCACTCCCAATCGGGGCTCCCGGAACTCTCTGGGCCATCC 1561  
Qy 1441 CAAGATAACGACCGGACTGGGGTATTGA 1470  
Db 1562 CAAGATAACGACCGGACTGGGGTATTGA 1591

## RESULT 5

US-10-377-718-3  
; Sequence 3, Application US/10377718  
; Publication No. US20030175291A1  
; GENERAL INFORMATION:  
; APPLICANT: KUO, Tsun Yuang  
; TITLE OF INVENTION: MULTIPLE AND MULTIVALENT DNA VACCINES IN OVO  
; FILE REFERENCE: 39734-186920  
; CURRENT APPLICATION NUMBER: US/10/377,718  
; CURRENT FILING DATE: 2003-03-04  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 15186  
; TYPE: DNA  
; ORGANISM: Newcastle disease virus (NDV)  
US-10-377-718-3

Query Match 78.1%; Score 1148.4; DB 15; Length 15186;  
Best Local Similarity 86.3%; Pred. No. 0;  
Matches 1269; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

Qy 1 ATGCTTCTCGTATTTCGATGAATACGAGCAGCTCCTCGTGTCTCAGACTCGGCCCAATGGA 60  
Db 122 ATGCTTCTCGTATTTCGATGAATACGAGCAGCTCCTCGTGTCTCAGACTCGGCCCAATGGA 181  
Qy 61 GCTCA CGGAGGGGAGAGAGAGGAGCAGCTTTAAGAGTTGAGGTCCTCAGTATTCACCTCTT 120  
Db 182 GCTCATGAGGGGAGAGAGAGGAGTACCTTTAAGTAGAGCTCCCGGTATTCACCTCTT 241  
Qy 121 AACAGTACGATTCAGAGATAGATGGAATTTTGGGTATTCGTTTCGGATTCGTT 180  
Db 242 AACAGTACGATTCAGAGATAGATGGAATTTTGGGTATTCGTTTCGGATTCGTT 301  
Qy 181 AGCAGGAGCGCAACAAACCGCTCAGGCAAGGTGCTCTCATATCCCTCTGTGCTCCCAT 240  
Db 302 AGCAGGATGCCAAACCACTCAGGCAAGGTGCTCTCATATCTCTTTATGCTCCCAT 361  
Qy 241 TCTCAAGTATGAGGAACCATGTTGCCCTTCAGGAAACAGAAATGAGGCTACACTGACT 300  
Db 362 TCACAGGTAAATGAGGAACCATGTTGCCATTCAGGAAACAGAAATGAGGCTACACTGACT 421  
Qy 301 GTTCTTGAATCGATGGTTTTCAGAGAGCTGCTCAGTTCAACAAACAGAGTGGGTG 360  
Db 422 GTGCTTGAATGGTGGTTTTCGCAACCGCACGCCCTCAGTTCAACAAATAGGAGTGGGTG 481  
Qy 361 TCTGAGGAGAGACAGAGATTCATGTTGATAGCGGCTCTCTCCCTCGGCGTGCAGT 420  
Db 482 TCTGAGGAGAGACAGAGATTCATGTTGATAGCGGCTCTCTCCCTCGGCGTGCAGT 541  
Qy 421 AACGCTACTCGTTCGTCACGCGCTGGGTTTCAAGATGATGCAACAGAAAGATATCACTCAT 480  
Db 542 AACGGAACCCGTTTCGTCACAGCGCGGCGAGAGTATGATGCAACAGAGATCACTCAT 601  
Qy 481 ACTCTGGAAGATTCCTGTTCTATCAGGCTCAGGTATGGGTCAAGTATGCAAGGCGCATG 540  
Db 602 ACCCTGGAGAGGATCTCTCTATCAGGCTCAAGTATGGGTCAAGTATGCAAGGCGCATG 661  
Qy 541 ACTCATATGAGACAGAGATGATTCGGAACAGAGAGATCAATAGTATGATATGAGCAA 600  
Db 662 ACTCGGTATGAGATGTCAGAGTATGCGAAACAGAGGAGATCAATAAGTATATGCGCAA 721  
Qy 601 GGCAAGTCCAGAAAGATGATCATCTCCACCTCTATGCGAGTGCATATTCATCTACA 660  
Db 722 GGCAGGCTCCAAAGAAATATCATCTCTACCCCTATGCGAGGACACATCACTCAG 781  
Qy 661 ATCAGACATTCCTGGCAGTCCGCAATTTCTTATGAGGAGCTTAAGAGAGGCGCAAT 720  
Db 782 ATCAGACAGTCTCTTGGAGTCCGCACTTTTGGTTAGCGAGCTCAAGAGAGGCGCAAC 841  
Qy 721 ACGGAGGTGGGAGCTCCAGTATTACAACTTATGAGGAGTGTAGACTCATACATCAGG 780  
Db 842 ACGGAGGTGGTACCTCTACTTATTATACCTGTTAGGGAAGTATGATCATCATCAGG 901  
Qy 781 AACACCGGACTTACTGCAATCTCTCTTACACTCAAAATATGGAATTAATACCAAGACATCA 840  
Db 902 AATACCGGCTTACTGCACTTCTTCTTGACACTCAAGTACGGAATCAACCAAGACATCA 961  
Qy 841 GCCCTAGCACTCAGCAGCTCTCAGGCGATATCCAAAGATGAAGCAGCTCATGCGTTTA 900  
Db 962 GCCCTTGCACCTTAGTAGCTCTCAGGCGACATCCAGAGATGAAGCAGCTCATGCGTTG 1021  
Qy 901 TATCGGATGAAGGAGAAATCGCGGTACATGACATTTGTTGCCATGGGATGGCATCAGATG 960  
Db 1022 TATCGGATGAAGGAGAAATCGCGGTACATGACATTTACTTGTGATGATGAGCCAGATG 1081  
Qy 961 AGCTTTGACCGGCTGAGTATGCACAGCTTTTATCTTTTGCCATGGGATGGCATCAGTC 1020  
Db 1082 AGCTTTGCGCTCGCGAGTATGCACAACTTTACTCTTTTGCCATGGGATGGCATCAGTC 1141

QY 1021 TTAGATAAGGAATCGGCAATACCAATTCGCCAGAGACTTCATGAGCACATCATCTTGG 1080  
DB 1142 CTAGATAAAGGTACTGGGAAATACCAATTTGCCAGGACTTTATGAGCACATCATCTTGG 1201  
QY 1081 AGACTCGGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAAAGACATGCTGCT 1140  
DB 1202 AGACTTGGAGTAGATGCTCAGGCTCAGGGAGTAGCATTAACGAGGATATGCTGCC 1261  
QY 1141 GAGCTAAACTAACCCCGGAGCAAGAAAGGGGCTTGGCAGAGTGTGCTGCCCAAGAGTGTCT 1200  
DB 1262 GAGCTAAAGCTAACCCAGCAGCAATGAAGGCTTGGCAGTGTGCTGCCCAAGGCTCTCC 1321  
QY 1201 GAGGAACTGGCAGCGTGGATATCTCTCAACGAGCGGGGCTCTCACTGGGCTCAGC 1260  
DB 1322 GACGATACGAGCAGATATACATGCTCTCAACAGTGGAGTCTCTACTGGGCTTAGC 1381  
QY 1261 GATGAGGCGCCCGGAGCCTCTCAGGGTGGATCGAAACAGTTCGCAAGGGGCAACAGATGCC 1320  
DB 1382 GAGGGGGGTCCCAAGCTCTCAAGGCGGATCGAATAGATCGCAAGGGGCAACAGAGCC 1441  
QY 1321 GGAGATGGGAGACCCCAATCTTTGGATTTGATGAGAGCAGTGGCGAAACAGCATGCGAGNA 1380  
DB 1442 GGGGATGGGAGACCCCAATCTTGGATCTGATGAGAGCGGTAGCAAAATAGCATGAGGAG 1501  
QY 1381 GCGCCAACTCGCAGAGACCCACCCAGCGGAAACCCCGGACTCCCGGGCCATCA 1440  
DB 1502 GGGCCAACTCTGCAAGGGGCACTCCCAATCGGGGCTCCCGGCACTCTCTGGGCCATCC 1561  
QY 1441 CAAGATAACGACACCGACTGGGGGTATTGA 1470  
DB 1562 CAAGATAACGACACCGACTGGGGGTATTGA 1591

RESULT 6

US-10-429-735-3  
; Sequence 3, Application US/10429735  
; Publication No. US20030207836A1  
; GENERAL INFORMATION:  
; APPLICANT: KUO, Tsun Yuang  
; TITLE OF INVENTION: VACCINE ACCELERATOR FACTOR (VAF) FOR IMPROVEMENT OF VACCINATIONS  
; FILE OF INVENTION: IN POULTRY  
; FILE REFERENCE: 39734-188449  
; CURRENT APPLICATION NUMBER: US/10/429, 735  
; CURRENT FILING DATE: 2003-05-06  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 3  
; LENGTH: 15186  
; TYPE: DNA  
; ORGANISM: Newcastle disease virus (NDV)  
US-10-429-735-3

Query Match 78.1%; Score 1148.4; DB 16; Length 15186;  
Best Local Similarity 86.3%; Pred. No. 0;  
Matches 1269; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 1 ATGCTCTCGTATTCGATGATACGAGCAGCTCTCGTCTCAGACTCGCCCAATGGA 60  
DB 122 ATGCTCTCGTATTTGATGATGATGAGCTCTCGGCTCAGACTCGCCCAATGGA 181  
QY 61 GCTCAGGAGGGGAGAGAGGAGGAGCATTAAAGAGTTGAGGTCCAGATTCACTCTT 120  
DB 182 GCTCAGGAGGGGAGAGAGGAGGATACCTTAAAGTAGACGTCGCGGTATTCACTCTT 241  
QY 121 AACAGTACGATCCAGAGATAGATGGAATTTTCGGGTATTCGTCTTCGGATTCGTGT 180  
DB 242 AACAGTATGACCCAGAGATAGATGAGCTTTTGGGTATTCGTCTTCGGATTCGTGT 301  
QY 181 AGCGAGAGCGCAACAAACGCTCAGGAGAGGTCTCTCATATCCCTCCTGTGCTCCCAT 240  
DB 302 AGCGAGATGCCAAACAAACCACTCAGGAGAGGTCTCTCATATCTCTTTATGCTCCAC 361  
QY 241 TCTCAAGTATGAGGAACCACTGTTGCCCTTTCAGGAGAAACAGATGAGGCTACACTGACT 300

DB 362 TCACAGGTAAATGAGGAACCAATGTTGCCATTCAGGAGAAACAGAAATGAAGCCACATTTGGCC 421  
QY 301 GTTCTTGAGATCGATGGTTTTACAGCAGCGCTCAGTTTCAACAAACAGGAGTGGGGTG 360  
DB 422 GTGCTTGAGATGATGGCTTTGCCAACCGGACGCCCGAGTTCAACAAATAGGAGTGGAGTG 481  
QY 361 TCTGAGGAGAGACACAGAGATTCATGCTGATAGCAGGGTCTCTCCCTCGGCGCTGCGAGT 420  
DB 482 TCTGAGAGAGAGACACAGAGATTTGCGATGATAGCAGGATCTCTCCCTCGGSCATGCGAGC 541  
QY 421 AACGCTACTCCGTTCTGTCACGCGTGGGTGGAAGATGATGACACAGAGATATCACTGAT 480  
DB 542 AACGGAACCCCGTTCGTCACAGCGCGGAGAGATGATGACCAAGAGACATCAACCGAT 601  
QY 481 ACTCTGGAAGAAATCTCTCTATCCAGGCTCAGGTATGCTCAGAGTACGAGAGCCCATG 540  
DB 602 ACCCTGAGAGAGATCTCTCTATCCAGGCTCAAGTATGGGTACAGATGCAAAAGCCATG 661  
QY 541 ACTGCATATGACACAGCAGATGAGTCGGAACCAAGAAAGATCAATAAGTACATGTCAGCAA 600  
DB 662 ACTGCTATGAGACTGCGAGATGAGTTCGGAACCAAGCGCAATCAATAAGTATATGAGCAA 721  
QY 601 GGCAGAGTCCAGAAAGATGATCTCTCCACCTGATGAGGAGTGCATTTCAACTCACA 660  
DB 722 GGCAGGCTCCAAAGAAATATACATCTCTACCCCGTATGCGAGGAGCAATCACTCAACTCAG 781  
QY 661 ATCAGACATCTCTCGGAGTCCGCAATTTCTTAGTTAGCGACTTAAAGAGAGCCGCAAT 720  
DB 782 ATCAGACAGTCTCTTGGAGTCCGCAATCTTTTGGTTAGCGACTCAAGAGAGCCGCAAC 841  
QY 721 AGCGCAGGTGGGAGGCTCCACGTTATTAACAATTTAGTAGGGGATGTAGACTCATATCAGG 780  
DB 842 AGCGCAGGTGGTACCTCTACTTATTAACCTGGTAGGGGACGTAGACTCATACATCAG 901  
QY 781 AACACCGGACTTACTGCAATCTCTTCACTCAATATGGAATTAATACCAAGACATCA 840  
DB 902 AATACCGGCTTACTGCAATCTCTTGAACACTCAAGTACGGAATCAACCAAGACATCA 961  
QY 841 GCCTTAGCACTCAGCAGGCTCAGCGGATATCCAAAGATGAAGCAGCTCATGCGTTTA 900  
DB 962 GCCTTTGCACTTAGTAGCTCTCAGCGGACATCCAGAGATGAGCAGCTCATGCGTTTG 1021  
QY 901 TATCGGATCAAGGGAGAAAATGCGCGTACATGACATTCAGGTCAAGTATGATGATG 960  
DB 1022 TATCGGATCAAGGAGATAATGCGCGTACATGACATTTCTTGGTATGATGACCAAGT 1081  
QY 961 AGCTTTGCAACCGGCTGAGTATGCAAGCTTTATCTTTTGGCATGGCATGGCATGATC 1020  
DB 1082 AGCTTTGCGCTGCGGAGTATGCACAACTTTACTCTCTTTGCCATGGGTATGGCATGATC 1141  
QY 1021 TTAGATAAAGGAATCGGCAAAATACCAATTCGCCAGAGACTTCATGAGCACATCATCTTGG 1080  
DB 1142 CTAGATAAAGGTACTGGGAAATACCAATTTGCCAGGACTTTATGAGCACATCATCTTGG 1201  
QY 1081 AGACTCGGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCATCAACGAAAGACATGCTGCT 1140  
DB 1202 AGACTTGGAGTAGATGCTCAGGCTCAGGGAGTAGCATTAACGAGGATATGCTGCTGCC 1261  
QY 1141 GAGCTAAACTAAACCCCGGAGCAAGAGGGGCTTGGCAGTGTGCTGCCCAAGAGTGTCT 1200  
DB 1262 GAGCTAAAGCTAAACCCCGAGCAATGAAGGGGCTTGGCAGTGTGCTGCCCAAGGCTCTCC 1321  
QY 1201 GAGGAACTGGCAGCGTGGATATTCCTACTCAACAGAGCCGGGCTCTCTACTGGGCTCAGC 1260  
DB 1322 GAGATACGAGCAGATATATACGCTCTCAACAGTCCGAGTCTCTACTGGGCTTAGC 1381  
QY 1261 GATGAGGCGCCCGGAGCCTCTCAGGGTGGATGGAACAAAGTTCGAAAGGGGCAACAGATGCC 1320  
DB 1382 GAGGGGGGTCCCAAGCTCTCAAGGCGGATCGAATAGATCGCAAGGGGCAACCAAGAGCC 1441  
QY 1321 GGAGATGGGAGAGACCCCAATCTTTGGATTTGATGAGCAGTGGCGAAACAGATGCGAGAA 1380

Db 1442 GGGGATGGGAGACCCAAATTCCTGGATCTGATGAGAGCGGTAGCAAAATAGCATGAGGGAG 1501  
 Qy 1381 GCGCAAACTCGCAGACAGACACCCACCGGAACCCCGGACTCCCGGGCCATCA 1440  
 Db 1502 GCGCAAACTCTGCACAGGCACTCCCAATCGGGGCTTCCCCCAACTCCTGGGCCATCC 1561  
 Qy 1441 CAAGATAACGACACCGACTGGGGGTATTGA 1470  
 Db 1562 CAAGATAACGACACCGACTGGGGGTATTGA 1591

## RESULT 7

US-09-951-061A-125  
 ; Sequence 125, Application US/09951061A  
 ; Publication No. US2003008204A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Paolletti, Enzo  
 ; APPLICANT: Tartaglia, James  
 ; APPLICANT: Taylor, Jill  
 ; APPLICANT: Gettig, Russell  
 ; TITLE OF INVENTION: POXVIRUS - CANINE DISTEMPER VIRUS (CDV)  
 ; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE  
 ; TITLE OF INVENTION: RECOMBINANTS  
 ; NUMBER OF SEQUENCES: 143  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell, Boehnen, Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/951.061A  
 ; FILING DATE: 13-SEP-2001  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 09/354,138  
 ; FILING DATE: 15-JUL-1999  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/224,657  
 ; FILING DATE: 16-APR-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/073,962  
 ; FILING DATE: 08-JUN-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/776,867  
 ; FILING DATE: 23-OCT-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/621,614  
 ; FILING DATE: 30-NOV-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/938,283  
 ; FILING DATE: 31-AUG-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/621,614  
 ; FILING DATE: 30-NOV-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/105,483  
 ; FILING DATE: 12-AUG-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/847,951  
 ; FILING DATE: 06-MAR-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/713,967  
 ; FILING DATE: 11-JUN-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07,666,056  
 ; FILING DATE: 07-MAR-1991  
 ; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.  
 ; REGISTRATION NUMBER: 25,506  
 ; REFERENCE/DOCKET NUMBER: 454310-2860  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 840-3333  
 ; TELEFAX: (212) 840-0712  
 ; INFORMATION FOR SEQ ID NO: 125:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1572 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; US-09-951-061A-125

Query Match 7.3%; Score 106.6; DB 10; Length 1572;  
 Best Local Similarity 46.0%; Pred. No. 1.6e-24;  
 Matches 361; Conservative 0; Mismatches 424; Indels 0; Gaps 0;  
 Qy 499 TCTATCCAGGCTCAGGTATGGGTACAGTAGGGAAGGCCATGACTGCATATGAGACGCA 558  
 Db 505 TCCATCTTTGGCTCAAATTTGGATCCTGCTAGCTAAAGCGGTGACTGCTCTTGATATGCA 564  
 Qy 559 GATGAGTCGGAAACAAAGAAGATCAATAAGTACATGCGAGCAAGGACAGAGTCCAGNAGAAG 618  
 Db 565 GCCGACTCGGAGATGAGAAGTGGATTAAAGTATACCCAGCAAGACGCTGTGGTCGAGAA 624  
 Qy 619 TACATCTCTCCACCTGTATGCGAGGAGTGCATTAACAATCACAATCAGACATTTCTTGGCA 678  
 Db 625 TTTAGAATGAACAAAATCTGGCTTGTATATTGTTAGAAACAGGATTTGTGAGGACTATCT 684  
 Qy 679 GTCCGCAATTTCTTAGTTAGCGAGCTTAAGAGAGGCGCAATAGCGGAGTGGAGCTCC 738  
 Db 685 TTGAGGCGATTTCATGGTGGCGCTCATCTTGGACATCAAAACGATCCCGAGAAACAAGCCT 744  
 Qy 739 ACGTATTACAACTTAGTAGGGGATGTAGACTCATACATCAGGAACAACGGACTTACTGCA 798  
 Db 745 AGAATTCGTGAATGATTTGTGATATAGATTAACATACATTTGGAAGCTGGGTTAGCTAGT 804  
 Qy 799 TTCTTCTTACACTCAAATATGGAATTAATACCAAGACATCAGCCCTAGCACTCAGCAGC 858  
 Db 805 TTCATCTTAATCAAGTTTGGCATTTGAACTATGTATCGGCTCTTGGGTTGCATGAG 864  
 Qy 859 CTACAGGCGATATCCAAAGATGAAGCAGCTCATGCGTTTATATCGGATGAAGGAGAA 918  
 Db 865 TTTTCGGGAGAAATTAACAACTATTGAATCCCTCATGATGCTTATATCAACAGATGGGTGA 924  
 Qy 919 AATGCGCGGTACATGACATTTGCTAGGTGACAGTATCAGATGAGCTTTGCAACCGGCTGAG 978  
 Db 925 ACAGCACCCTACATGGTTATCTTGGAAAACTCTGTTCAAAACAAATTTAGTGCAGGGTCC 984  
 Qy 979 TATGCAAGCTTTTATCTTTTGGCATGGGATGGCATCAGTCTTAGATAAAGAACTGGC 1038  
 Db 985 TACCCATTGCTCTGGAGTTATGCTATGGGGTTGGTGTGAACTTGAATCTCATGGGA 1044  
 Qy 1039 AATATCAATTTCCGAGAGACTTCATGAGCAGATCATTTCTGGAGACTCGGGGTGGAGTAT 1098  
 Db 1045 GGGTTAAATTTCCGTGCTTACTTTGACCCAGCTTACTTTCAGACTCGGCGCAAGAAATG 1104  
 Qy 1099 GCTCAGGCTCAGGGGAGTAGCATCAACGAGACATGCTGCTGCTGAGCTTAAACTAACCCCG 1158  
 Db 1105 GTTAGAGATCTCCCGGCAAGTAAGCTCTGCACTTTCGCGCCGAGCTTGGCATCACCAAG 1164  
 Qy 1159 GCAGCAAGAGGGGCGCTGGCAGCTGCTGCCAACGAGTGTCTGAGAAATCTGGCAGCGTG 1218  
 Db 1165 GAGGAAGCTCAGCTAGTGTGAGAAATAGCATCAAGACACACAGAGACCGGACAAATCGA 1224  
 Qy 1219 GATATTCCTACTCAACAGCGGGGTCTCTCACTGGGCTCAGCGATGAGGCCCCCGAGCC 1278  
 Db 1225 GCTACTGGTCTTAAGCAATCCCAATCACTTTTCTGCACTCGGAAAGATCCGAAAGTCGCC 1284  
 Qy 1279 TCTCA 1283

```

; LENGTH: 1578 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-951-061A-134

Query Match          6.8%; Score 10.2; DB 10; Length 1578;
Best Local Similarity 45.6%; Pred. No. 2.3e-22;
Matches 351; Conservative 0; Mismatches 418; Indels 0; Gaps 0;

QY 500 CTATCCAGGCTCAGGTATGGGTCA CAGTAGCGAAGCCATGACTGCAATGAGACGACG 559
Db 506 CCATCTCTAGCCAAAATTTGGGCTTTGCTCGAAAAGCGGTTACGGCCCGACAGACGGCAG 565
QY 560 ATGAGTCGGAACACAGAAGAAACAATTAAGTACTGATCGACGAAGGAGAGTCCAGAAAGAGT 619
Db 566 CTGATTCGGAGCTAAGAGAGGTGATTAAGTACACCCACAAAGAAGGTGATTTGGTGAAT 625
QY 620 ACATCTCTCCACCCTGTATCGAGGAGTGCAATTCAACTCAATCFAGACATTTCTCTGGCAG 679
Db 626 TTAGATTGGAGAGAAAATGGTTGGATGTGCTGAGGAACAGGATTTGCCGAGGACCTCTCTCT 685
QY 680 TCCGCAATTTCTTAGTTAGGAGCTTAAAGAGGCGCGCAATACGGCAGGTGGGAGCTCCA 739
Db 686 TAGCGCGAATTCATGGTCGCTCTAATCCTGGATATCAAGAGAACACCCGGGAAACAAACCCA 745
QY 740 CGPATTACAACCTTAGTAGGGGATGTAGACTCATACATCAGGAACACCGGACTTACTGCAAT 799
Db 746 GGATTGCTGAATGATATGTGACATTTGATACATATATCGTAGGCGAGGATTAGCCAGTT 805
QY 800 TCTTCTCTTACACTCAAATATGGAATTAATACCAAGACATCAGCCCTAGCACTCAGCAGCC 859
Db 806 TTATCTCTGACTATAAGGTTTGGGATAGAAACTATGTATCTCTGCTCTTGGACTGCATGAAT 865
QY 860 TCACAGCGGATATCCAAAAGATCAAGCAGCTCATCGTTTATATCGATGAAGGAGAGAAA 919
Db 866 TTGCTGTGTAGTTATCCACACTTGAGTCTTGTGATGAACCTTTACAGAGCAATGGGGGAAA 925
QY 920 ATGCGCGGTACATCACAATTTGCTAGGTGCACAGTGCATCAGATAGCTTTGCAACCGGTGAGT 979
Db 926 CTGCACCTACATGTTGTAATCTCTGGAGAACTCAATTTCAGAACAGTTTCAGTGCAGGATCAT 985
QY 980 ATGCAAGCTTTTATCTTTTGCCATGGGCATGSCATCAGTCTTAGATTAAGGAACCTGGCA 1039
Db 986 ACCCTCTGCTCTGGAGCTATGCCATGGGAGTAGGAGTGGAACTTTGAAACTCCATCGGGG 1045
QY 1040 AATACCAATTTGCGCAGAGACTTCATGAGCACATCATTTCTGGAGACTCGGGGTGGAGTATG 1099
Db 1046 GTTTGAATCTTGGCCGATCTTACTTTGATCCAGCATATTTTAGATTAGGCAAGAGATGG 1105
QY 1100 CTCAGGCTCAGGGAGTAGCATCAACGAAGACATGGCTGTCTGAGCTAAACTTAACCCCG 1159
Db 1106 TAAGGAGGTGAGCTGGAAAGCTCAGTTTCCATCTGGCATCTGAACCTCGGTATCACTGCCG 1165
QY 1160 CAGCAAGAGGGGCTCGCAGCTGTCTGCCCAACGAGTGTCTGAGGAAACTGGCAGAGTGG 1219
Db 1166 AGGATGCAAGGCTGTGTTTCAGAGATTGCAATGCATCTACTGAGGACAAATCAGTAGAG 1225
QY 1220 ATATTCTTACTCAACAGCCGGGTCTCTCACTGGGCTCAGCGATGGAGG 1268
Db 1226 CGGTTGGACCCAGACAAGCCCAAGTATCATTTTCTACACGGTGTCAAG 1274

RESULT 9
US-09-900-112-35
; Sequence 35, Application US/09900112
; Publication No. US20030082209A1
; GENERAL INFORMATION:
; APPLICANT: Skiadopoulos, Mario H.
; APPLICANT: Collins, Peter L.
; APPLICANT: Murphy, Brian R.
; APPLICANT: Schmidt, Alexander C.

```

Db 1285 AATCA 1289

RESULT 8  
US-09-951-061A-134  
; Sequence 134, Application US/09951061A  
; Publication No. US20030082204A1  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: Tartaglia, James  
; APPLICANT: Taylor, Jill  
; APPLICANT: Gettig, Russell  
; TITLE OF INVENTION: POXVIRUS - CANINE DISTEMPER VIRUS (CDV)  
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE  
; TITLE OF INVENTION: RECOMBINANTS  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell, Boehnen, Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/951,061A  
; FILING DATE: 13-SEP-2001  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 09/354,138  
; FILING DATE: 15-JUL-1999  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/224,657  
; FILING DATE: 16-APR-1994  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/073,962  
; FILING DATE: 08-JUN-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/776,867  
; FILING DATE: 23-OCT-1991  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/621,614  
; FILING DATE: 30-NOV-1990  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/938,283  
; FILING DATE: 31-AUG-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/621,614  
; FILING DATE: 30-NOV-1990  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/105,483  
; FILING DATE: 12-AUG-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847,951  
; FILING DATE: 06-MAR-1992  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/713,967  
; FILING DATE: 11-JUN-1991  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07,666,056  
; FILING DATE: 07-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2860  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 134:  
; SEQUENCE CHARACTERISTICS:



; TITLE OF INVENTION: Attenuated Human-Bovine Chimeric Parainfluenza Virus (PIV) Vaccin
; FILE REFERENCE: NIH-0127
; CURRENT APPLICATION NUMBER: US/09/900,112
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/215,809
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 15456
; TYPE: DNA
; ORGANISM: bovine PIV3 Ka strain
US-09-900-112-35

Query Match 5.2%; Score 76; DB 10; Length 15456;
Best Local Similarity 54.7%; Pred. No. 1.5e-13;
Matches 151; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 755 TAGGGGATGTAGACTCATACATCAGGAACACCGGACTTACTGCAATCTTCTTACACTCA 814
DB 868 TACAGATTGTAGGAACACTACATCAGAGATGCGAGTCTTGCTTCATTTTCAACAAATCA 927
QY 815 AATATGGAATTAATACCAACATCAGCCCTAGCACTCAGCAGCCTCAGAGGGGATATCC 874
DB 928 GATATGCAATGTAGACTAGAGTGGCAGCTCTAACTCTGTCTACCTTAGACCGGATATCA 987
QY 875 AAGAATGAAGCAGCTCATCGTTTATATCGGATGAAGGAGAGAAATGCGCCGCTACATGA 934
DB 988 ACAGACTCAAGGCACATGATCGAGTTATATCTATCAAAAGGGGCCAGCGTCTCTTTATAT 1047
QY 935 CANTGTAGTGTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 994
DB 1048 GCATTTGAGAGATCCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
QY 995 CTTTGGCATGGGATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGG 1030
DB 1108 GTTATGCGATGGTGTAGCAGTTGTACAAAACAAGG 1143

RESULT 10
US-09-900-112-36
; Sequence 36, Application US/09900112
; Publication No. US20030082209A1
; GENERAL INFORMATION:
; APPLICANT: Skiadopoulos, Mario H.
; APPLICANT: Collins, Peter L.
; APPLICANT: Murphy, Brian R.
; APPLICANT: Schmidt, Alexander C.
; TITLE OF INVENTION: Attenuated Human-Bovine Chimeric Parainfluenza Virus (PIV) Vaccin
; FILE REFERENCE: NIH-0127
; CURRENT APPLICATION NUMBER: US/09/900,112
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/215,809
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 15456
; TYPE: DNA
; ORGANISM: bovine PIV3 SF strain
US-09-900-112-36

Query Match 5.2%; Score 76; DB 10; Length 15456;
Best Local Similarity 54.7%; Pred. No. 1.5e-13;
Matches 151; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 755 TAGGGGATGTAGACTCATACATCAGGAACACCGGACTTACTGCAATCTTCTTACACTCA 814
DB 868 TACAGATTGTAGGAACACTACATCAGAGATGCGAGTCTTGCTTCATTTTCAACAAATCA 927
QY 815 AATATGGAATTAATACCAACATCAGCCCTAGCACTCAGCAGCCTCAGAGGGGATATCC 874
DB 928 GATATGCAATGTAGACTAGAGTGGCAGCTCTAACTCTGTCTACCTTAGACCGGATATCA 987

875 ARAAGATGAGCAGCTCATCGTTTATATCGGATGAGGAGAGAAATGCGCCCTACATGA 934
988 ACAGACTCAAGGCACATGATCGAGTTATATCTATCAAAAGGGGCCAGCGTCTCTTTATAT 1047
935 CATTGTAGTGTACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 994
1048 GCATTTTGTAGAGATCCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
995 CTTTGGCATGGGATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGG 1030
1108 GTTATGCGATGGTGTAGCAGTTGTACAAAACAAGG 1143

RESULT 11
US-10-302-547-131
; Sequence 131, Application US/10302547
; Publication No. US20040142448A1
; GENERAL INFORMATION:
; APPLICANT: MURPHY, BRIAN R.
; APPLICANT: COLLINS, PETER L.
; APPLICANT: SKIADOPOULOS, MARIO H.
; APPLICANT: NEWMAN, JASON T.
; TITLE OF INVENTION: RECOVERY OF RECOMBINANT HUMAN PARAINFLUENZA VIRUS TYPE 1 (HPIV1) FROM cDNA AND USE OF RECOMBINANT HPIV1 IN IMMUNOGENIC COMPOSITIONS AND AS VECTORS TO ELICIT IMMUNE RESPONSES AGAINST PIV AND OTHER HUMAN PATHOGENS
; FILE REFERENCE: 2303-37-3
; CURRENT APPLICATION NUMBER: US/10/302,547
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 60/331,961
; PRIOR FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 131
; LENGTH: 15609
; TYPE: DNA
; ORGANISM: Human parainfluenza virus 1
US-10-302-547-131

Query Match 5.0%; Score 73.2; DB 17; Length 15609;
Best Local Similarity 48.6%; Pred. No. 1.4e-12;
Matches 201; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

QY 762 TGTAGACTCATACATCAGGACACCGGACTTACTGCAATCTTCTTACACTCAAAATGG 821
DB 887 TGTAGGAAATTCATAAGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 946
QY 822 AATTATATCAACATCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 881
DB 947 TGTAGAAACGAGATGCGCCCTGACACTATCAATCTGAGCAGGATATACAAATTT 1006
QY 882 GAAGCAGCTCATCGTTTATATCGGATGAAGGAGAGAAATGCGCGGTACATGATGCT 941
DB 1007 GAGAAAGCCTTGTGATATCTATCTATCAAAAGGAGCGGAGCGCCCTTTTATGATGACT 1066
QY 942 AGTGACAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1001
DB 1067 CAGAGACCCAGTTTCATGGAGACTTTCCTTGGAACTATCCAGCACTGTGGAGCTACGC 1126
QY 1002 CATGGGATGGCATCGCTTGTAGTAAAGGAACTGGCAAAATACCAATTCGCGCAGAGACTT 1061
DB 1127 AATGGGCTTGTGTGTGATCAAAACAAAGCTATGCAACAGTATGTAAGTGGAGACATA 1186
QY 1062 CATGAGCAGTATCTTGTGAGACTCGGGGTGGAGTATGCTCAGGCTCAGGGGAGTAGCAT 1121
DB 1187 TTTGGACATGAAATGTTCTTACTTGGACAAGCTGTAGCTTAAAGATGCTGATTTCCAAAT 1246
QY 1122 CAAGAGACATGCTGCTGAGCTAAACTAACCCCGGAGCAGAGAGGGGCT 1175
DB 1247 CAGCAGTGTCTTGGAGGAGAACTAGGTGTGACAGATACAGCAAAAGAGAGACT 1300

## RESULT 12

US-09-733-692A-60  
; Sequence 60, Application US/09733692A  
; Patent No. US20020155581A1  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Brian R.  
; APPLICANT: Collins, Peter L.  
; APPLICANT: Schmidt, Alexander C.  
; APPLICANT: Durbin, Anna P.  
; APPLICANT: Skiadopoulos, Mario H.  
; APPLICANT: Tao, Tao  
; TITLE OF INVENTION: USE OF RECOMBINANT PARAINFLUENZA VIRUSES (PIVs) AS  
; TITLE OF INVENTION: VECTORS TO PROTECT AGAINST INFECTION AND DISEASE CAUSED  
; TITLE OF INVENTION: BY PIV AND OTHER HUMAN PATHOGENS  
; FILE REFERENCE: 15280-404100US  
; CURRENT APPLICATION NUMBER: US/09/733,692A  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 60  
; LENGTH: 15492  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Sequence of  
; OTHER INFORMATION: pFLC.PIV32, 15492 bp in sense orientation  
US-09-733-692A-60

Query Match 4.9%; Score 72.2; DB 9; Length 15492;  
Best Local Similarity 54.3%; Pred. No. 3e-12;  
Matches 146; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY	762	TGTAGACTCATACATCAGAACACCGGACTTACTGCACTTCTTCTTACACTCAAAATATGG	821
Db	875	TGTTGGCAACTACATAAGAGATGCAGGTCTCGCTTCATCTTCAATACATCAGATATGG	934
QY	822	AATTAATACCAAGATCAGCCCTAGCACTCAGCGGCTCAGCGGATATCCAAAGAT	881
Db	935	AATTGAGCCAGAAATGGCAGCTTTGACTCTATCCACTCTCAGACCATATCAATAGATT	994
QY	882	GAAGCAGCTCATCGCTTTATATCGATGAAGGGAGAAATGCGCGTACATGACATTGCT	941
Db	995	AAAAGCTTTGATGGAACGTATTTATCAAGGGACCCAGCGCTCTTCTATCTGTATCCT	1054
QY	942	AGGTGACAGTCAATGCGTTTATATCGATGAAGGGAGAAATGCGCGTACATGACATTGCT	941
Db	1055	CAGAGATCTTATACATGATGAGTTTCGACACGAGCAACTATCTCTGCCATATGGAGCTATGC	1114
QY	1002	CATGGGATGGCATCAGTCTTAGATAAAG	1030
Db	1115	AATGGGGTGGCAGTTGTACAAATAGAG	1143

## RESULT 13

US-09-733-692A-62  
; Sequence 62, Application US/09733692A  
; Patent No. US20020155581A1  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Brian R.  
; APPLICANT: Collins, Peter L.  
; APPLICANT: Schmidt, Alexander C.  
; APPLICANT: Durbin, Anna P.  
; APPLICANT: Skiadopoulos, Mario H.  
; APPLICANT: Tao, Tao  
; TITLE OF INVENTION: USE OF RECOMBINANT PARAINFLUENZA VIRUSES (PIVs) AS  
; TITLE OF INVENTION: VECTORS TO PROTECT AGAINST INFECTION AND DISEASE CAUSED  
; TITLE OF INVENTION: BY PIV AND OTHER HUMAN PATHOGENS  
; FILE REFERENCE: 15280-404100US  
; CURRENT APPLICATION NUMBER: US/09/733,692A  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 15498  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Sequence of  
; OTHER INFORMATION: pFLC.PIV32TW, 15498 bp in sense orientation  
US-09-733-692A-61

Query Match 4.9%; Score 72.2; DB 9; Length 15498;  
Best Local Similarity 54.3%; Pred. No. 3e-12;  
Matches 146; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY	762	TGTAGACTCATACATCAGAACACCGGACTTACTGCACTTCTTCTTACACTCAAAATATGG	821
----	-----	--------------------------------------------------------------	-----

; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 62  
; LENGTH: 15492  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Sequence of  
; OTHER INFORMATION: pFLC.PIV32CT, 15474 bp in sense orientation  
US-09-733-692A-62

Query Match 4.9%; Score 72.2; DB 9; Length 15492;  
Best Local Similarity 54.3%; Pred. No. 3e-12;  
Matches 146; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY	762	TGTAGACTCATACATCAGAACACCGGACTTACTGCACTTCTTCTTACACTCAAAATATGG	821
Db	875	TGTTGGCAACTACATAAGAGATGCAGGTCTCGCTTCATCTTCAATACATCAGATATGG	934
QY	822	AATTAATACCAAGATCAGCCCTAGCACTCAGCGGCTCAGCGGATATCCAAAGAT	881
Db	935	AATTGAGCCAGAAATGGCAGCTTTGACTCTATCCACTCTCAGACCATATCAATAGATT	994
QY	882	GAAGCAGCTCATCGCTTTATATCGATGAAGGGAGAAATGCGCGTACATGACATTGCT	941
Db	995	AAAAGCTTTGATGGAACGTATTTATCAAGGGACCCAGCGCTCTTCTATCTGTATCCT	1054
QY	942	AGGTGACAGTCAATGCGTTTATATCGATGAAGGGAGAAATGCGCGTACATGACATTGCT	941
Db	1055	CAGAGATCTTATACATGATGAGTTTCGACACGAGCAACTATCTCTGCCATATGGAGCTATGC	1114
QY	1002	CATGGGATGGCATCAGTCTTAGATAAAG	1030
Db	1115	AATGGGGTGGCAGTTGTACAAATAGAG	1143

## RESULT 14

US-09-733-692A-61  
; Sequence 61, Application US/09733692A  
; Patent No. US20020155581A1  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Brian R.  
; APPLICANT: Collins, Peter L.  
; APPLICANT: Schmidt, Alexander C.  
; APPLICANT: Durbin, Anna P.  
; APPLICANT: Skiadopoulos, Mario H.  
; APPLICANT: Tao, Tao  
; TITLE OF INVENTION: USE OF RECOMBINANT PARAINFLUENZA VIRUSES (PIVs) AS  
; TITLE OF INVENTION: VECTORS TO PROTECT AGAINST INFECTION AND DISEASE CAUSED  
; TITLE OF INVENTION: BY PIV AND OTHER HUMAN PATHOGENS  
; FILE REFERENCE: 15280-404100US  
; CURRENT APPLICATION NUMBER: US/09/733,692A  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 15498  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Sequence of  
; OTHER INFORMATION: pFLC.PIV32TW, 15498 bp in sense orientation  
US-09-733-692A-61

Query Match 4.9%; Score 72.2; DB 9; Length 15498;  
Best Local Similarity 54.3%; Pred. No. 3e-12;  
Matches 146; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY	762	TGTAGACTCATACATCAGAACACCGGACTTACTGCACTTCTTCTTACACTCAAAATATGG	821
----	-----	--------------------------------------------------------------	-----



```
Db      875 TGTGGCACTACATAAGAGATGAGGTCTCGCTTCATTCTTCAATACATCAGATATGG 934
Qy      822 AATTATATCCAAAGACATCAGCCCTAGCACTCAGCAGCCTCAGCGGATATCCAAAAGAT 881
      |||||
Db      935 AATTGAGACCAAGATGGCAGCTTTGACTCTATCCACTCTCAGACCAGATATCAATAGATT 994
      |||||
Qy      882 GAAGCAGCTCATGCGTTTATATCGATCAAGGGAGAAATGCGCGGTACATGACATTGCT 941
      |||||
Db      995 AAAAGCTTTGATGGAACGTGTATTTATCAAGGGACCGCGCTCCTTTCACTGTATCCT 1054
      |||||
Qy      942 AGGTGACAGTGATCAGATGAGCTTTTGACCCGGCTGAGTATGCACAGCTTTATTTCTTTTC 1001
      |||||
Db      1055 CAGAGATCCATACATGCTGAGTTGCGACCAAGGCAACTATCTCGCATATGGAGCTATGC 1114
      |||||
Qy      1002 CATGGGCATGGCATCAGCTCTTAGATAAAG 1030
      |||||
Db      1115 AATGGGGTGGCAGTTGTACAAAATAGAG 1143
      |||||
```

## RESULT 15

```
US-09-764-891-10172
; Sequence 10172, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10172
; LENGTH: 11809
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-10172
```

```
Query Match      2.7%; Score 39.8; DB 10; Length 11809;
Best Local Similarity 60.7%; Pred. No. 0.25;
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy      1333 ACCCAATCTTTGGATTTGATGAGAGCAGTGGGGAACAGCATGCGAGAGCGCCAACTCC 1392
      |||||
Db      5643 ATCCAGTCAGGGAGATCCTGAGGATGGGAGCGGACAGACTGGAACTCCAGGCTCCAGGAGG 5702
      |||||

Qy      1393 GCACAGAGCACACCCACCCCGAAACCCCGGACTCCCGGGGCCATC 1439
      |||||
Db      5703 GCAAAGAACCCACCCAGCTGTGAACCCCGCCCGCCCGCCCGCCAGC 5749
      |||||
```

Search completed: September 4, 2004, 00:41:31  
Job time : 1004 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 3, 2004, 19:36:35 ; Search time 5764 Seconds  
(without alignment)  
7615.789 Million cell updates/sec

Title: US-09-970-851-1  
Perfect score: 1470  
Sequence: 1 atgtcttcctgattcattgatga.....acacgactgggggtattga 1470

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estnu:\*
  - 5: em\_estov:\*
  - 6: em\_estpl:\*
  - 7: em\_estro:\*
  - 8: em\_htc:\*
  - 9: gb\_est1:\*
  - 10: gb\_est2:\*
  - 11: gb\_est3:\*
  - 12: gb\_est4:\*
  - 13: gb\_est5:\*
  - 14: gb\_est6:\*
  - 15: em\_estfun:\*
  - 16: em\_estom:\*
  - 17: em\_gss\_hum:\*
  - 18: em\_gss\_inv:\*
  - 19: em\_gss\_pln:\*
  - 20: em\_gss\_vrt:\*
  - 21: em\_gss\_fun:\*
  - 22: em\_gss\_nam:\*
  - 23: em\_gss\_mus:\*
  - 24: em\_gss\_pro:\*
  - 25: em\_gss\_rod:\*
  - 26: em\_gss\_phg:\*
  - 27: em\_gss\_vrl:\*
  - 28: gb\_gss1:\*
  - 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	48	3.3	910	29	CNS0060N
C 2	46.4	3.2	1201	13	BX356664
C 3	41.6	2.8	1201	13	BX442599
C 4	41.4	2.8	839	29	CNS004NB
C 1	48	3.3	910	29	CNS0060N
C 2	46.4	3.2	1201	13	BX356664
C 3	41.6	2.8	1201	13	BX442599
C 4	41.4	2.8	839	29	CNS004NB

C 5	41.2	2.8	1201	13	BX381961
C 6	41	2.8	452	10	BB785240
C 7	40.6	2.8	940	28	CC421824
C 8	40.2	2.7	475	9	AL450622
C 9	40.2	2.7	835	13	BX468739
C 10	39.6	2.7	486	9	AI917768
C 11	39.6	2.7	1201	13	BX360624
C 12	39.6	2.7	1208	29	AG115851
C 13	39.4	2.7	895	10	BE571023
C 14	39.4	2.7	988	29	CNS0079R
C 15	39.4	2.7	1201	13	BX406481
C 16	38.8	2.6	849	13	BX462111
C 17	38.8	2.6	885	13	BX425603
C 18	38.6	2.6	298	10	BB026608
C 19	38.4	2.6	600	12	B1813308
C 20	38.4	2.6	1101	29	CNS00F3N
C 21	38.2	2.6	344	14	CF244892
C 22	38.2	2.6	975	29	CC837496
C 23	38.2	2.6	1042	28	CC246998
C 24	38	2.6	1260	13	BQ677411
C 25	37.8	2.6	1101	29	CNS016HG
C 26	37.6	2.6	459	29	CE329457
C 27	37.6	2.6	991	28	CC241937
C 28	37.4	2.5	419	9	AI009304
C 29	37.2	2.5	494	14	CE713434
C 30	37.2	2.5	551	14	CD660629
C 31	37.2	2.5	640	14	CD466270
C 32	37.2	2.5	651	14	CD535076
C 33	37.2	2.5	828	29	CNS048X6
C 34	37.2	2.5	993	13	BX400687
C 35	37.2	2.5	1064	29	CNS0204S
C 36	37	2.5	454	14	CD407576
C 37	37	2.5	510	12	BG905028
C 38	36.8	2.5	435	12	B1437139
C 39	36.8	2.5	633	13	BU546640
C 40	36.8	2.5	682	28	AQ656834
C 41	36.8	2.5	716	28	BH519387
C 42	36.8	2.5	1201	13	BX401671
C 43	36.8	2.5	1287	10	BF664989
C 44	36.8	2.5	2319	13	BU199108
C 45	36.8	2.5	3183	11	BC034944

ALIGNMENTS

RESULT 1  
CNS0060N/c

LOCUS  
DEFINITION

CNS0060N 910 bp DNA linear GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence T7 end of BAC #  
BACR14J21 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION  
VERSION

AL065629  
AL065629.1 GI:4944698

KEYWORDS  
SOURCE

Drosophila melanogaster (fruit fly)  
Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 910)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

REFERENCE  
AUTHORS

TITLE  
JOURNAL

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooeagawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of



```
/db_xref="taxon:9606"
/clone="CS0DF033YF22"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 2.8%; Score 41.6; DB 13; Length 1201;
Best Local Similarity 35.9%; Pred. No. 1.8; Indels 0; Gaps 0;
Matches 97; Conservative 40; Mismatches 133; Indels 0; Gaps 0;

QY 1111 GCGAGTAGCATCAACGAAGACATGGCTGTGAGCTAAACCTAACCCCGGAGCAAGG 1170
DB 1066 SGGANAMCCGDKCCSACCRCMCCGGAAGACRAAARCCAGKCGKGMWMC 1007

QY 1171 GGCCTGGAGCTGTCGCCAAGAGTGTCTGAGGAACTGGCAGCGTGATATTCCTACT 1230
DB 1006 SKCDAGTGGKGTCCMAAAAGACYGKGTGAACCCSKAAACMSKGRITSMCCCCC 947

QY 1231 CAACAGCGGGTCTCTACTGGCTCAGCATGGAGGCCCGCGAGCTCTCAGGGTGA 1290
DB 946 CRCCCGGGGGYCCCCGAGAGGTGACCCSCAAGGAGAGTGAGMAGGGRGMAAGC 887

QY 1291 TCGAACAACTGCGAAGCGCAACAGATGCGGAGATGGGAGACCCAAATCTTGGATTG 1350
DB 886 CAGAAAAGGGAGGGGGCCMCWAGCCGGBGACSCGGGTCCCGCAAKCAKCMAG 827

QY 1351 ATGAGAGCATGGCGCAACAGCATGCGAGAA 1380
DB 826 AGGAGKNCACAAAGGAGGRRGTGGAGAA 797

RESULT 4
CNS004NB 839 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR10E16 of RPI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL054280 1 GI:4931788
GSS.
Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 839)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers

FEATURES
source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1072YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
```

```
source
1. .839
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR10E16"
/clone_lib="RPI-98"
/note="end : TET3"
```

## ORIGIN

```
Query Match 2.8%; Score 41.4; DB 29; Length 839;
Best Local Similarity 19.5%; Pred. No. 1.7;
Matches 60; Conservative 113; Mismatches 134; Indels 0; Gaps 0;

QY 1128 AGCATGGCTGTGAGCTAAACCTAACCCCGGAGCAAGGCTGTGAGCTGTGC 1187
DB 474 ASAGSGGCGASCAACSSACAGSAGSSASGSCAGSSSSAGSGGASGSRSG 533

QY 1188 CCAACAGTGTCTGAGGAACTGGCAGCGTGATATTCCTACTCAACAGCCGGGTCT 1247
DB 534 GAGSGGGGGGSRGCGGAGVGSAGSSSACSCSASSSSSSSSSASGCMSCSSSSSAA 593

QY 1248 CACTGGGCTCAGGATGGAGGCCCCGAGCCTCTCAGGTGGATCGAACAGTCGCAAG 1307
DB 594 SCSSSSSSGVAACVCSGSGSGVGVSAAAVAVASGARMGMGAGGAVSAAASAA 653

QY 1308 GCAACAGATGCGGAGATGGGAGACCCAAATCTTGGATTGATGAGACAGTGGCGAA 1367
DB 654 ASAAAVGAAACSSGSGAGSSAGAAARACASVCASSSSGSCGSSSMCSAVSSGAA 713

QY 1368 CAGCATGCGAGAGCGCCAACTCCGACAGAGCACACCCACCCGGAACCCCCCGAC 1427
DB 714 SSGASGCGCGSSSGSCSCSSSSSCSSSGGSSVCSGSSVCSGSSSCSBSBSSC 773

QY 1428 TCCCGG 1434
DB 774 SCCASS 780
```

## RESULT 5

```
QY 1428 TCCCGG 1434
DB 774 SCCASS 780

RESULT 5
BX381961/c 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1072YF05 3-PRIME, mRNA sequence.
ACCESSION BX381961
VERSION BX381961.1 GI:30453007
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1072CC03NP1.
Location/Qualifiers
```

## FEATURES

```
source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1072YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
```

sites of the pCMVSPORT 6 vector. Library was normalized."

[illegible]

RESULT 6	BB785240	452 bp	mrna	linear	EST 08-JUL-2003
BE785240/c	BB785240	RIKEN full-length enriched,	embryo RCB-0549	Cle-H3	cdna
LOCUS	Mus musculus	cDNA clone G430101B20 3',	mRNA sequence.		
DEFINITION	BB785240				
ACCESSION	BB785240.1	GI:16953736			
VERSION	EST.				
KEYWORDS	Mus musculus	(house mouse)			
SOURCE					

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1. (bases 1 to 452)

REFERENCE  
AUTHORS

Akimura, T., Arakawa, T., Hiramoto, K., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, R., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.

Journal Comment

Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-5222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: <http://genome.gsc.riken.go.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Mateuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA Encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
source

```

Location/Qualifiers
1. .452
/organism="Mus musc
/mol_type="mRNA"
/db_xref="taxon:100
/clone="G430101B20"
/cell_line="RCB-054
/dev_stage="embryo"
/clone_lib="RIKEN f
cle-H3_cdna"

```

## ORIGIN

Query Match	2.8%; Score 41; DB 10; Length 452;
Best Local Similarity	51.4%; Pred. No. 1.6;
Matches	95; Conservative 0; Mismatches 90; Indels 0; Gaps 0
QY	450 TGAATCATATCACAGAAGATATCACTGATACTCTGGAAAGAAATCCTGTCTATCCAGC 509
Db	361 TGAATATTGGACAGAGAAACCAGGGCTACAGGGGAAAAAATAATCACCATGATGA 302
QY	510 TCAGTATGGGTCAAGTAGCGAAGGCCATGACTGTCATATGACACAGCAGATGAGTCGA 569
Db	301 TCAAAACAGAGAAACGAAAGGAAATATGTCAATATCAGGAAGACACGCTGAAACAAAACC 242
QY	570 AACAGAGAGATCAATAGTATACATGCAGCAAGCCAGAGTCCAGAAAGATCATCTCTCA 629
Db	241 AACAAACCCATCAGTAGGACATTTATGAAGCAAAATAATGAAGGTGTACATACTAAA 182
QY	630 CCTTG 634
Db	181 CACAG 177

RESULT 7  
CC421824/C

CC421824	940 bp	DNA	linear	GSS 19-MAY-2003
PUHEA37TD	ZM 0.6_1.0 KB	mayas	genomic clone	ZMMBTa430S01,
DEFINITION	genomic survey sequence.			

ACCESSION CC421824  
 VERSION CC421824.1  
 KEYWORDS GSS.

**SOURCE**  
**ORGANISM**

REFERENCE

1 (bases 1 to 940)

clade; Panicoidae; Andropogoneae; Zea.

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE  
AUTHORS

**TITLE**  
JOURNAL  
COMMENT

Benitz, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennett, J.  
Maize Genomics Consortium  
Unpublished (2003)  
Other GSSs: PUHEA37TB  
Contact: Cathy Whitelaw  
TIGR

TEL: 301-838-5843  
Fax: 301-838-0208  
Email: [whitelaw@tigr.org](mailto:whitelaw@tigr.org)  
Seq primer: TF

Class: sheared ends.

# FEATURES

Location/Qualifiers  
1. 940  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZM0430301"  
/clone\_lib="ZM 0.6-1.0 kb"  
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
Cor selected genomic DNA library"

## ORIGIN

Query Match 2.8%; Score 40.6; DB 28; Length 940;  
Best Local Similarity 53.8%; Pred. No. 3.2; Mismatches 89; Indels 1; Gaps 1;  
Matches 105; Conservative 0;  
QY 685 ATTTCTTTAGGAGCTTAAGAGAGCGCCCAATACGGCAGGTGGGAGCTCCACGTAT 744  
Db 223 ATACTATTTTCAAGGAACCTAAGAGAGAAAATAAATAAAGTAGAAAAATAATAC 164  
QY 745 TACAACCTAGTAGGGATGAGACTACATCAGGAACACCGGACTTACTGCAATCTTC 804  
Db 163 AATAA-TTCCAAATGAATGTAAACATCTATTGGAAAGAGATGAATGATTGAATTTCTT 105  
QY 805 CTTACACTCAATATGGAATTAATACCAAGACATCAGCCTAGCACTCAGCAGCTTACA 864  
Db 104 GATCATTAAGAGAGGTATAGTTTCAAAACATCAGTCATCCAGTCAGAGTTCTCAAG 45  
QY 865 GCGCATATCCAAAG 879  
Db 44 GGAACACCCCAAG 30

## RESULT 8

AL450622 475 bp mRNA linear EST 07-DEC-2000  
LOCUS subsp. vulgare Hordeum vulgare subsp. vulgare  
DEFINITION subsp. vulgare cDNA clone HK03H20u 5', mRNA sequence.  
ACCESSION AL450622  
VERSION AL450622.1 GI:11602030  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Hordeum.  
1 (bases 1 to 475)  
REFERENCE Michalek W., Weschke, W., Pleissner, K.-P. and Graner, A.  
EST sequencing and analysis in barley  
Unpublished (2000)  
Contact: Michalek W  
Institute for Plant Genetics and Crop Plant Research  
Corrensstr. 3, D-06466 Gatersleben, Germany  
Email: michalek@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de  
Seq primer: M13uni primer for 5' end.

# FEATURES

Location/Qualifiers  
1. 475  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="barke"  
/sub\_species="vulgare"  
/db\_xref="taxon:112509"  
/clone="HK03H20u"  
/tissue\_type="etiolated leaves"  
/lab\_host="Xl1Blue"  
/clone\_lib="Hordeum vulgare Barke etiolated leaves"  
/note="Vector: pBluescript SK; Site 1: SalI; Site 2: NotI; mRNA was made from etiolated leaves of spring barley variety 'Barke', a high quality malting variety. Plants were grown on filterpaper for 6 d at 25 C in the dark. Leaves were harvested at daylight conditions. Sal-Adaptor Sequence: TCGACCCACGCTCG Average insert size is 1 kb

Sequence trimming: Vector sequences and sequence ends were trimmed from the 5'- and 3'-end until a 50 bp window contains less than two ambiguities. The maximum length was set to 700 bp."

## ORIGIN

Query Match 2.7%; Score 40.2; DB 9; Length 475;  
Best Local Similarity 55.3%; Pred. No. 2.8; Mismatches 63; Indels 0; Gaps 0;  
Matches 78; Conservative 0;  
QY 1241 GGGTCTCTCACTGGGCTAGCGATGAGGCGCCCGAGCTCTCAGGTTGGATCGAACAGT 1300  
Db 74 GGTTCGAGCTCTGTCGCCGACATGGCGTGCCTGCGCTCAAGTGGACCGGAGGCG 133  
QY 1301 CGCAAGGCAACACAGATCCGGAGATGGGAGAGACCAATTCTTGGATTGGATGAGAGCAG 1360  
Db 134 GGCACCTTCGTGGAGCGCGGAGCTGGAGCGGCGGTGAGTGCTCTGATGGCGGAGGAG 193  
QY 1361 TGGCGACAGCATGCGAGAAG 1381  
Db 194 GGGAGGACGCCCGGAGGAG 214

## RESULT 9

BX468739 835 bp mRNA linear EST 12-JUN-2003  
LOCUS subsp. vulgare Anopheles gambiae cDNA clone NAPI-P138-H-07-5, mRNA  
DEFINITION subsp. vulgare Anopheles gambiae cDNA clone NAPI-P138-H-07-5, mRNA  
ACCESSION BX468739  
VERSION BX468739.1 GI:31659680  
KEYWORDS EST.  
SOURCE Anopheles gambiae (African malaria mosquito)  
ORGANISM Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.  
1 (bases 1 to 835)  
REFERENCE Christophides, G.K., Blass, K., Zdobnov, E.M., Carmouche, R., Benes, V. and Kafatos, F.C.  
Unpublished (2002)  
TITLE Anopheles gambiae EST, European Molecular Biology Laboratory  
JOURNAL Contact: Christophides GK  
COMMENT Foris C. Kafatos laboratory  
European Molecular Biology Laboratory  
Meyerohofstrasse 1, 69117 Heidelberg, Germany  
Tel: +49 6221 387-440  
Fax: +49 6221 387-306  
Email: christop@embl-heidelberg.de  
Contact: Christophides G.K.  
European Molecular Biology Laboratory  
Meyerohofstr. 1, 69117 Heidelberg, Germany.  
Tel: +49 6221 387-440  
Fax: +49 6221 387-306  
Email: christop@embl-heidelberg.de  
Plate: P138 row: H column: 07.  
FEATURES  
Location/Qualifiers  
1. 835  
/organism="Anopheles gambiae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7165"  
/clone="NAPI-P138-H-07-5"  
/lab\_host="E. coli DH10B"  
/clone\_lib="NAPI"  
/note="Vector: pT73D-Pac (Pharmacia); Site 1: NotI; Site 2: EcoRI; ESTs sequenced from the T7 priming site that reads from the 5' end of cDNA. The NAPI is a directionally cloned and normalized, oligo-T primed cDNA library constructed from a mixture of Anopheles gambiae developmental stages according to: Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery, Genome Research 6, 791-805."

## FEATURES

Location/Qualifiers  
1. 835  
/organism="Anopheles gambiae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7165"  
/clone="NAPI-P138-H-07-5"  
/lab\_host="E. coli DH10B"  
/clone\_lib="NAPI"  
/note="Vector: pT73D-Pac (Pharmacia); Site 1: NotI; Site 2: EcoRI; ESTs sequenced from the T7 priming site that reads from the 5' end of cDNA. The NAPI is a directionally cloned and normalized, oligo-T primed cDNA library constructed from a mixture of Anopheles gambiae developmental stages according to: Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery, Genome Research 6, 791-805."

## ORIGIN



Query Match 2.7%; Score 40.2; DB 13; Length 835;  
 Best Local Similarity 54.4%; Pred. No. 3.9;  
 Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 291 TACACTGACTGTTCTTGGAGATCGATGGTGTATACAGCAGCGTCTCAGTTCAACAACAG 350  
 DB 180 TAACTGGATGTAATTGATATCGAGTTTGGCTGTCCCGTGTCCAGTTTCCAGTATTCACAC 239

QY 351 GAGTGGGTGTTCTGAGGAGAGACAGAGATTGATGATAGCAGGTCTCTCCCTCG 410  
 DB 240 GGGGGGAATTTGTTTGCATGTTGACAGAAATGCTTTAATGCGGAAGGCGCGAAGCCGG 299

QY 411 GCGGTGCACTACGGTACTCCGTTGCTCA 439  
 DB 300 CTGGGAATTTACGGGACAACGTTGGTTA 328

RESULT 10  
 AI917768/C  
 LOCUS AI917768 486 bp mRNA linear EST 17-DEC-1999  
 DEFINITION wbl5904.x1 NCI\_CGAP\_GC6 Homo sapiens cDNA clone IMAGE:2305782 3',  
 mRNA sequence.  
 ACCESSION AI917768  
 VERSION AI917768  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 486)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
 R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 1052 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 462.

FEATURES  
 source  
 1..486  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2305782"  
 /tissue\_type="pooled germ cell tumors"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP GC6"  
 /notes="Vector: p773D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA  
 from the normalized library NCI\_CGAP GC4 was prepared, and  
 ss circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (clonids  
 1257096-1258631, 1469064-1470983, and 1475592-1476743).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 2.7%; Score 39.6; DB 9; Length 486;  
 Best Local Similarity 47.9%; Pred. No. 4.3;  
 Matches 114; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 248 TGATGAGGAACCATGTTGCTTGCAGGAAACAGATGAGGTACTGACTGTCTTG 307  
 DB 381 TGTGAGGAACCAATGTTGCTTCTGTCATCAACAGCCACCTCTCTATCTCTAA 322

QY 308 AGATCGATGGTTTACAGCAGCGTCTCAGTTCAACAACAGGAGTGGGTGCTGAGG 367  
 DB 321 ATGATCTGGCTTTTCTTGGAGGCTGCCATGTTTGAAGATGGTATCAGAGGGCTGCT 262

QY 368 AGAGACACAGAGATTATGTTGATAGCAGGCTCTCTCCCTGGCGCTGCAGTAACGGTA 427  
 DB 261 GGGCAGTCTGCTTCGGGCGCAGGCTCAGGACCCCTCTGCTCTGCGAGCTTAACTGTC 202

QY 428 CTCCCTTGGTCAACGCTGGGGTGAAGATGATGACACGAGAGATATCACTGATCTCT 485  
 DB 201 CTCTGCTAGGACCAAGGTTTCAAGCCAGGAGCAACTGGGACCCCTGAAAACTGT 144

RESULT 11  
 BX360624  
 LOCUS BX360624 1201 bp mRNA linear EST 05-MAY-2003  
 DEFINITION clone CS01072YL05 3-PRIME, mRNA sequence.  
 ACCESSION BX360624  
 VERSION BX360624  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1201)  
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 4982.r  
 Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS01072CF03NP1.

FEATURES  
 source  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS01072YL05"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dt)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 2.7%; Score 39.6; DB 13; Length 1201;  
 Best Local Similarity 11.6%; Pred. No. 7.4;  
 Matches 34; Conservative 152; Mismatches 105; Indels 1; Gaps 1;

QY 1173 CTGTCAGCTGTGTCACACGAGTGTCTGAGGAACTGCGAGGTGATTCCTACTCA 1232  
 DB 904 SSSSSSSSTAGSWMASASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 963

QY 1233 ACAAGCCGGGTCTCTACTGGCTCAGCGTACGAGGCCCGCCGAGCTCAGGGTGCATC 1292  
 DB 964 SSSSSSSSBSSBSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1023

QY 1293 GAACAGTCGCAAGGGCAACCGATCCCGGAGATGGGAGACCCCAATCTTGGATTGAT 1352  
 DB 1024 SVSVVASSRAASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 1083

QY 1353 GAGAGCAGTGGCGAACAGCATGCGGAGAGCGCCCAAACTCGGCACAGACACCCACC 1412



